UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 09/674,277 | 02/13/2001 | Dominique Therese Marie Frechon | P66034US0 | 5117 |

136 7590 06/21/2005

JACOBSON HOLMAN PLLC
400 SEVENTH STREET N.W.
SUITE 600
WASHINGTON, DC 20004

| EXAMINER | |
|---|---|
| DUFFY, PATRICIA ANN | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 1645 | |

DATE MAILED: 06/21/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 09/674,277 | FRECHON ET AL. | |
| | Examiner | Art Unit | |
| | Patricia A. Duffy | 1645 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *27 December 2004*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *20-25 and 27-60* is/are pending in the application.

4a) Of the above claim(s) *31-60* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *20-25 and 27-30* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☒ Other: *sequence attachments*.

## RESPONSE TO AMENDMENT

The amendments to the claims filed on 5-12-04 and 12-27-04 have been entered into the record. The amendment to the specification filed on 7-12-04 has been entered into the record. The responses filed 5-12-04, 7-12-04 and 12-27-04 have been entered into the record. Claims 1-19 and 26 have been cancelled. Claims 20-25, and 27-60 are pending.

The text of Title 35 of the U.S. Code not reiterated herein can be found in the previous office action.

### *Election/Restrictions*

Upon reconsideration of the lack of unity, and the fact that the individual SEQ ID NOS: 1 and 2 were examined in the first office action on the merits, Groups I -III are hereby rejoined. The lack of unity with respect to the corresponding method claims and primer pairs is maintained in view that these lack unity of invention in light of the art maintained. The traversal is on the ground(s) that the cited reference no longer anticipates claim 20 and as such, the claims now have unity of invention. This is not found persuasive because the deletion of "insertion" does not remove the art because "mutation" or "substitution" as it is still recited in the claims, broadly encompasses substitutions, insertions and deletions. Substitutions base X could be X for X+1, 2, 3... resulting in a substitution of a single base for multiple replacements or substitution of bases XYZ could be substituted with XZ resulting in the loss of a single or multiple bases and the "consisting of language" does not limit the claims to shorter fragments because a nucleic acid that is larger can also specifically detect and hybridize. Therefore, claims 20-25, 27-30 are under examination. Claims 31-60 are withdrawn to inventions that lack unity of invention for reasons made of record.

The requirement is still deemed proper and is therefore made FINAL.

***Rejections Withdrawn***

The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

***Rejections Maintained***

Claims 20 and 21 stand rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention is maintained for reasons made of record for claim 1.

The amendment to the claims does not moot the rejection. As previously made of record, *Vas-Cath Inc. v. Mahurkar*, 19 USPQ2d 1111, 1117, makes clear that "applicant must convey with reasonable clarity to those skilled in the art that, *as of the filing date sought*, he or she was in possession *of the claimed invention*". "The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed* [emphasis added]". Therefore, Applicants' arguments are not persuasive. SEQ ID NO: 1 and 2 fragments of the p0157 plasmid of the prior art. No written description is provided in the specification for any other species of nucleic acids that are derived therefrom by mutation, deletion and/or substitution with the claimed function of specifically detecting enterohemorrhagic E. coli. The disclosure of a these discrete sequences (*which the claims are not limited toward*) does not reasonably constitute the claimed genus of nucleic hybridizing nucleic acids that are mutated, deleted or substituted in one or more bases. Analogous to the situation decided in *Fiers v. Revel*, 25 USPQ2d 1601, 1606 (Fed. Cir. 1993), "an adequate written description of a DNA [product] requires more than a mere statement that it is part of the invention and reference to a potential method for isolating it; what is required is a description of the DNA itself". *Fiddes v. Baird*, 30 USPQ2d 1481, 1483 (1993) held that claims directed to mammalian FGFs were found unpatentable due to lack of written

description for the broad class, in which the specification had provided an adequate description of only the bovine sequence. Accordingly, the court held in *Univ. California v. Eli Lilly and Co.*, 43 USPQ2d 1398 (Fed. Cir. 1997) that: "One skilled in the art therefore cannot, as one can do with a fully described genus, visualize or recognize the identity of the members of the genus. A definition by function, as we have previously indicated, does not suffice to define the genus because it is only an indication of what the gene does, rather than what it is" and that: "A description of a genus of cDNAs [products] may be achieved by means of a recitation of a representative number of cDNAs [products], *defined by nucleotide sequence*, failing in the scope of the genus or of a recitation of structural features common to the members of the genus, *which features constitute a substantial portion of the genus* [emphasis added]. This is analogous to enablement of a genus under 112, [first paragraph], by showing the enablement of a representative number of species within the genus. See *Angstadt*, 537 F.2d at 502-03, 190 USPQ at 218". The specification does not teach any subsequence of either SEQ ID NO:1 or 2, that was derived by mutation, deletion and/or substitution of one or more bases. There is no written description for any sequences derived, much less those that would hybridize.

The rejection is maintained over the claimed sequences as "derived from SEQ ID NO:1 or 2 by mutation, deletion and/or substitution of one or more bases".

Claims 20, 21 and dependent claims 22-25 stand rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention is maintained for reasons made of record for the previous claims.

Applicants' arguments have been carefully considered. Applicants argue that representative examples of such high stringency conditions are recited in the specificaiton. This is not persuasive, exemplification is not a definition of specific conditions and limitations as recited in the specification are not read into the claims.

Further, the conditions relied upon at page 7, lines 19-30 are admittedly specifically related to the specific chemical structure of the hybridizing nucleic acid, the structure of which is not defined in the claims. The rejection is therefore maintained.

Claims 20, 21, 22 and 24 stand rejected under 35 U.S.C. 102(b) as being clearly anticipated by Brunder et al (Microbiology, 146:3305-3315, 1996) for reasons made of record in the Office Action Mailed 12-13-03.

Applicants' arguments have been carefully considered but are not persuasive. Applicants argue that the inventors have demonstrated that *E. coli* 0157:7 is characterized by the stable integration of a portion of the *IS91* sequence into the *katP* gene. According the stable combination of a portion of *IS91* with a portion of *kat P* is a specific marker for *E. coli* 0157:H7 strains. This is not persuasive, the claims are not so limited because "a portion" is a single "nucleotide" and the claims specifically encompass mutation, deletion or substitution of one or more bases or different ones and that the sequence of the prior art does not detect other EHEC's as set forth in the claims. The specific junction relied upon as set forth in the Figure is not defined in claim 21 and neither is the "portion" of sequence of *IS91* or "portion" of gene sequence of *katP*. Additionally, the "substitution could represent the substitution of "X" for "X + 12 or more additional bases". Further, plasmid pSm10 or pSM9 comprises the Sma-1 fragment of pO157 and contains more than 1 Kb upstream of the beginning of *kat P* (see page 3307, Figure 1) as derived from of would specifically hybridize to SEQ ID NO:1, because as with SEQ ID NO:1 is derived from the 0157 enterohemorrhagic plasmid of the prior art. Further, the *Sma-1* fragment inherently has the claimed nucleotide residues of claim 22. It is also noted that the first *HindIII* fragment of the *Sma-1* fragment of pSm10 or pSmK9 as disclosed by Brunder et al would also inherently hybridize under the high stringency conditions and would be specific for the sequence that they were derived from. The specifically claimed nucleotide sequence is inherent to the plasmids. The art is applied

against the nucleic acids of claims 24 and 25 because the claims are confusing as dependent from claims 20 and 21 respectively for reasons set forth in the second paragraph rejection above.

Claims 20, 21, 22, 24 and 25 are rejected under 35 U.S.C. 102(b) as being clearly anticipated by Makino et al , (DNA Research, 5(1):1-9, Feb 28, 1998) in light of GenEMBL Accession Number AB011549.

Makino et al teach the isolation of and complete nucleotide sequences of 93 kb and 3.3 kb plasmids of an enterohemorrhagic *Escherichia coli* 0157:H7 derived from Sakai outbreak. Makino et al teach extraction of DNA from the bacterium and isolation of the pO157 plasmid and subsequent sequencing of the plasmid, see page 2, Materials and Methods. Makino et al teach that pO157 is represented by EMBL Accession Number AB011549 (page 2, column 2, second full paragraph). The isolated plasmid inherently characterized by the stable integration of a portion of the *IS91* sequence into the *katP* gene and inherently hybridizes under the asserted conditions because it is 99.2% identical as compared to SEQ ID NO:1 (see attached alignment). As previously set forth, the language of mutation, deletion and/or substitution of one or more bases specifically includes the differences between SEQ ID NO:1 and the plasmid of the art. The deletion of "insertion" does not remove the art because "mutation" or "substitution" as it is still recited in the claims, broadly encompasses substitutions, insertions and deletions. Mutations include deletions and substitutions base X could be X for X+1, 2, 3 nts... resulting in a substitution of a single base for multiple replacements or substitution of bases XYZ could be substituted with XZ resulting in the loss or addition of a single or multiple bases and the "consisting of language" does not limit the claims to shorter fragments because a nucleic acid that is larger can also specifically detect and hybridize The isolated plasmid inherently hybridizes under stringent conditions because it comprises a sequence that is 98.7% identical as compared to SEQ ID NO:2 as evidenced by GenEMBL

Accession Number AB011549 (see attached alignment). Applicants argue that the sequence was not available prior to the priority date of this application. This is not persuasive, the filing date of the instant specificaiton in the US is the instant filing date and the filing date of the international application is 4-27-1999. Applicants the priority document of 99/05329 filed 4-28-98 is in a foreign language and such the priority is not perfected. Even if the foreign priority is perfected, the paper was available as of February 29, 1998 and fully enabled as of this date. The nucleic acid sequence is inherent to the isolated plasmid. Claimed residues 400-407 of SEQ ID NO:1, SEQ ID NOS: 10-13, 18-20, 21-23 and 25 are inherently contained in the hybridizing plasmid of the prior art (see attached alignments for SEQ ID NOS:1 and 2).

*New Rejections Based on Amendment*

The amendment filed 7-12-04 is objected to under 35 U.S.C. 132(a) because it introduces new matter into the disclosure. 35 U.S.C. 132(a) states that no amendment shall introduce new matter into the disclosure of the invention. The added material which is not supported by the original disclosure is as follows: Applicants indicate a correction to the specification based on a foreign patent document of French origin. It is noted where a foreign priority document under 35 U.S.C. 119 is of record in the U.S. application file, applicant may not rely on the disclosure of that document to support correction of an error in the pending U.S. application. *Ex parte Bondiou,* 132 USPQ 356 (Bd. App. 1961). This prohibition applies regardless of the language of the foreign priority documents because a claim for priority is simply a claim for the benefit of an earlier filing date for subject matter that is common to two or more applications, and does not serve to incorporate the content of the priority document in the application in which the claim for priority is made. This prohibition does not apply where the U.S. application explicitly incorporates the foreign priority document by reference. Applicants have not provided for a specific incorporation by references in the originally filed transmittal documents or

in the first line of the specification. As such, reliance on an incorporation by reference to correct the error is impermissible in this situation. Where a U.S. application as originally filed was in a non-English language and an English translation thereof was subsequently submitted pursuant to 37 CFR 1.52(d), if there is an error in the English translation, applicant may rely on the disclosure of the originally filed non-English language U.S. application to support correction of an error in the English translation document. In this case the originally filed document was not filed in a non-English language. Applicants are directed to MPEP section 2163.07 for correcting "obvious errors" not based on the priority documents.

Claims 20-25 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The is a new matter rejection.

The claims now require the nucleic acid to "specifically detect" enterohemorrhagic *E. coli* wherein it is a fragment or derived from SEQ ID NO:1 that specifically detects enterohemorrhagic *E. coli* (EHECs) wherein the fragment or derived sequence contains a nucleotide sequence of SEQ ID NO:1 resulting from a stable combination of at least a portion of insertion sequence IS91 and at least a portion of gene sequence KatP. It is noted that the specification teaches that this portion 400-407 is *specific to 0157:H7* and does not detect other EHEC's (see page 5, lines 12-20) as it relates to claims 21, 22, 24 (SEQ ID NOS:12, 13, 18, 19 and 20). As such, the use of this junction that is specifically described in the specification as unique to 0157:H7, as an attribute of ECECs in general is considered new matter because the specification specifically teaches that it is not a genus marker.

Claims 20-25 and 27-30 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

With respect to claims 20-25, the claims now recite that the isolated nucleic acid "specifically detect" enterohemorrhagic *Escherichia coli* (EHECs). In the absence of any definition in the specification, this phrase has been interpreted for this rejection as *exclusive detection of the target* (i.e. EHECs) to rule out detection of *all others*. The specification fails to disclose any individual/single isolated nucleic acid that specifically detects EHECs and some of the recited isolated nucleic acids are present in other genomes that are not EHEC's and an isolated nucleic acid consisting of SEQ ID NOS:10-13, 18-27 *per se* are of apparently insufficient length to form a stable hybrid under conditions such that they could *detect only EHECs* using single nucleic acid detection method such as hybridization. The specification is devoid of written description of specificity analysis with single nucleic acid probes for any of the particularly claimed nucleic acids. For example, with respect to specificity of detection using a single nucleic acid, residues 400-407 of SEQ ID NO:1 are present in fungi (see attached alignment) ; SEQ ID NO:19 is present in *Pseudomonas syringeae* DNA for IS801 insertion sequence (see attached alignment) and SEQ ID NO:18 is present in *Salmonella paratyphi* A (see attached alignment) and porcine liver factor XII (see attached alignment) and therefore these nucleic acids are not specific as claimed and one would have reason to doubt the asserted truth that the others as specifically claimed are able to "specifically detect" detect EHECs as claimed. In the absence of further guidance from Applicants as to which single nucleic acids "specifically detect EHECs" as compared to other microorganism under which hybridization or other conditions. The primer pairs do not support the specific detection using a single nucleic acid. Primer pairs amplify a fragment that is specific and they

themselves do not necessarily have to be exclusive. Therefore, the use of a specific primer pair to amplify a nucleic acid to generate a third nucleic acid sequence that has been identified as specific to 0157 is specifically distinguished from that which is claimed herein.

With respect to claims 27-30, Applicant's referral to the deposit of the clones pDF3 and pDF4 on page 5 of the specification is an insufficient assurance that all required deposits have been made and all the conditions of 37 CFR §1.801-1.809 have been met. If the deposit has been made under the provisions of the Budapest Treaty, filing of an affidavit or declaration by applicant or assignees or a statement by an attorney of record who has authority and control over the conditions of deposit over his or her signature and registration number stating that the deposit has been accepted by an International Depository Authority under the provisions of the Budapest Treaty, that all restrictions upon public access to the deposit will be irrevocably removed upon the grant of a patent on this application and that the deposit will be replaced if viable samples cannot be dispensed by the depository is required. This requirement is necessary when deposits are made under the provisions of the Budapest Treaty as the Treaty leaves this specific matter to the discretion of each State. *Amendment of the specification to recite the date of deposit and the complete name and full street address of the depository is required.* If the deposits *have not* been made under the provisions of the Budapest Treaty, then in order to certify that the deposits comply with the criteria set forth in 37 CFR §1.801-1.809, assurances regarding availability and permanency of deposits are required. Applicant's attention is directed to In re Lundack, 773 F.2d. 1216, 227 USPQ 90 (CAFC 1985) and 37 CFR §1.801-1.809 for further information concerning deposit practice.

Claims 20, 21 and dependent claims 22-25 stand rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

The claims now recite that the isolated nucleic acid "specifically detects" however, specificity is not defined in the specification nor is it defined as exclusive binding and therefore the metes and bounds of the claims can not be ascertained.

As to claims 22-25, the claims recite "the nucleic acid according to claim X" and it is not clear if the claims are intended to reference the "fragment" or "derived sequence". As such, it is not clear what alternative is specifically being limited in the dependent claims and therefore the dependent claims do not have proper antecedent basis or are properly dependent from independent claims 20 or 21.

As to claim 24, SEQ ID NOS:10 and 11, these sequences do not apparently have "at least a portion of IS91 and at least a portion of gene sequence katP, which appears to bridge residues 406-407 of SEQ ID NO:1 (see Figure 1). This issue is best resolved by Applicants pointing to the corresponding residues of each of these sequences that correspond to "the claimed "at least a portion of insertion sequence IS91" and "at least a portion of gene sequence katP".

Claims 20-25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Makino et al , (DNA Research, 5(1):1-9, Feb 28, 1998) in view of Schmidt et al, Microbiology 142(4):907-914, 1996 and Kennell et al 1971( "Principles and properties of nucleic acid hybridization", Progr. Nucl. Acid Res. Mol. Biol. 11: 259-301) in light of GenEMBL Accession Number AB011549 .

Makino et al is set forth *supra*. Makino et al differs by not teaching fragments of SEQ ID NO:1 or SEQ ID NO:2. Kennel et al teach the location of the open reading frames in the sequenced p0157 plasmid, including the hly operon. Makino et al identifies

the open reading frames of the junction of IS91/katP by nucleotide number and another open reading frame 82 corresponding to residues 87658-88761.

Schmidt et al teach that circular restriction fragment map of wild-type plasmid p0517 and location of the EHEC-hly operon (see page 910, Figure 2 and Table 1).

Kennell et al teach that nucleic acid sequences having a minimum size for stable complex formation is from 10-20 nucleotides depending upon the G+C content (see paragraph bridging pages 260-261).

It would have been *prima facie* obvious one of ordinary skill in the art at the time that the invention was made to use any restriction fragment of the p0157 plasmid of Schmidt et al of at least 10-20 nucleotides in length as a probe or primer to detect the p0157 of Makino et al or open reading fragment thereof because Makino et al teach that *E. coli* 0157:H7 is a pathogen that causes bloody diarrhea and hemorrhagic colitis and Kennel et al teach that the nucleic acids would form stable hybrids. To restate, any fragment of the p0157 plasmid from *E. coli* 0157:H7 having at least 10-20 nucleotides in length is *prima facie* obvious as a probe/primer for detecting the sequence of origin from enterohemorrhagic *E. coli* 0157:H7. The restriction fragment(s) would specifically detect the sequence from which it was derived. The art is applied against the nucleic acids of claims 24 and 25 because the claims are confusing as dependent from claims 20 and 21 respectively for reasons set forth in the second paragraph rejection above.

## *Status of Claims*

All claims stand rejected.

## *Conclusion*

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL.** See MPEP

§ 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Patricia A. Duffy whose telephone number is 571-272-0855. The examiner can generally be reached on M-Th 6:30 am - 6:00 pm. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Lynette Smith can be reached on 571-272-0864.

The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Patricia A. Duffy

Primary Examiner

Art Unit 1645

```
Qy       240 CGGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAGCGGGCTATTTCAG 299
Db      7076 CGGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAGCGGGCTATTTCAG 7135

Qy       300 GATACCCTTCGTCATCAACACGTACAAACCAGAAGACCAGCTTTTTGTTTCTGACATCCA 359
Db      7136 GATACCCTTCGTCATCAACACGTACAAACCAGAAGACCAGCTTTTTGTTTCTGACATCCA 7195

Qy       360 CAAAGAAGGGAATATTCAGGTCTGCGCAGCACTCAACGGCATCGTCAGTTGCGGCTTGGA 419
Db      7196 CAAAGAAGGGAATATTCAGGTCTGCGCAGCACTCAACGGCATCGTCAGTTGCGGCTTGGA 7255

Qy       420 ACCCCTTAGTATTTTTTGTCTGTAGTATCTATCCCAGCAATAGGTATATCCTGTTGCATC 479
Db      7256 ACCCCTTAGTATTTTTTGTCTGTAGTATCTATCCCAGCAATAGGTATATCCTGTTGCATC 7315

Qy       480 AATAAAGTTGACTTTTGTATACAACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATG 539
Db      7316 AATAAAGTTGACTTTTGTATACAACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATG 7375

Qy       540 ATAAAAAAAACTCTTCCTGTTCTGATTCTTCTGGCGCTATCGGGGAGCTTTTCTACCGCT 599
Db      7376 ATAAAAAAAACTCTTCCTGTTCTGATTCTTCTGGCGCTATCGGGGAGCTTTTCTACCGCT 7435

Qy       600 GTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT 659
Db      7436 GTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT 7495

Qy       660 CCTCTGAGATTACACAGCCCTGAATCAAATCCCTGGGGGGCTGATTTTGATTATGCCACC 719
Db      7496 CCTCTGAGATTACACAGCCCTGAATCAAATCCCTGGGGGGCTGATTTTGATTATGCCACC 7555

Qy       720 AGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAAGATTTGCTGACAACT 779
Db      7556 AGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAAGATTTGCTGACAACT 7615

Qy       780 TCCCAGGATTGGTGCCCTGCGGATTATGGTCATTATGGTCCTTTCTTTATTCGTATGGCT 839
Db      7616 TCCCAGGATTGGTGGCCTGCGGATTATGGTCATTATGGTCCTTTCTTTATTCGTATGGCT 7675

Qy       840 TGGCACGGTGCCGGAACATACAGGACATATGATGGCCGGGGAGGCGCCAGTGGTGGTCAG 899
Db      7676 TGGCACGGTGCCGGAACATACAGGACATATGATGGCCGGGGAGGCGCCAGTGGTGGTCAG 7735

Qy       900 CAACGTTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCGA 959
Db      7736 CAACGTTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCGA 7795

Qy       960 TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCCTGGGGAGACCTGATGGTC 1019
Db      7796 TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCCTGGGGAGACCTGATGGTC 7855

Qy      1020 CTGACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAAACGCTGGGATTTGCTGGCGGA 1079
Db      7856 CTGACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAAACGCTGGGATTTGCTGGCGGA 7915

Qy      1080 AGAGAAGATGACTGGGAGTCGGACCTGGTATACTGGGGGCCTGACAACAAGCCTCTTGCA 1139
Db      7916 AGAGAAGATGACTGGGAGTCGGACCTGGTATACTGGGGGCCTGACAACAAGCCTCTTGCA 7975

Qy      1140 GATAACCGGGATAAAAACGGGAAACTTCAGAAACCTCTTGCCGCCACGCAGATGGGACTT 1199
Db      7976 GATAACCGGGATAAAAACGGGAAACTTCAGAAACCTCTTGCCGCCACGCAGATGGGACTT 8035

Qy      1200 ATTTATGTCAATCCTGAAGGCCCCGGTGGAAAACCAGATCCTCTGGCTTCCGCGAAAGAT 1259
Db      8036 ATTTATGTCAATCCTGAAGGCCCCGGTGGAAAACCAGATCCTCTGGCTTCCGCGAAAGAT 8095

Qy      1260 ATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG 1319
Db      8096 ATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG 8155

Qy      1320 GGAGGGCATACATTTGGTAAAGCACATGGTGCAGCGTCTCCTGAAAAATGTATTGGCGCA 1379
Db      8156 GGAGGGCATACATTTGGTAAAGCACATGGTGCAGCGTCTCCTGAAAAATGTATTGGCGCA 8215

Qy      1380 GGGCCTGATGGTGCACCTGTGGAGGAGCAGGGACTGGGATGGAAAAATAAATGTGGTACA 1439
Db      8216 GGGCCTGATGGTGCACCTGTGGAGGAGCAGGGACTGGGATGGAAAAATAAATGTGGTACA 8275

Qy      1440 GGAAACGGCAAATATACCATCACCAGTGGCCTGGAAGGAGCCTGGTCGAC 1489
Db      8276 GGAAACGGCAAATATACCATCACCAGTGGCCTGGAAGGAGCCTGGTCGAC 8325
```

```
RESULT 5
AB011549
LOCUS       AB011549               92721 bp    DNA     circular BCT 27-APR-1999
DEFINITION  Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION   AB011549
VERSION     AB011549.2  GI:4589740
KEYWORDS    ToxR-regulated lipoprotein; tagA.
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (sites)
  AUTHORS   Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
            Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
            Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
            and Shinagawa,H.
  TITLE     Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
            enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
            outbreak
  JOURNAL   DNA Res. 5 (1), 1-9 (1998)
  MEDLINE   98290540
   PUBMED   9628576
REFERENCE   2  (bases 1 to 92721)
  AUTHORS   Makino,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1998) Kozo Makino, Research Institute for
            Microbial Diseases, Osaka University, Molecular Microbiology;
            Yamadaoka, 3-1, Suita, Osaka 562, Japan
            (E-mail:makino@bkns01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
            Fax:81-6-879-8320)
COMMENT     On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="O157:H7"
                     /sub_strain="RIMD 0509952"
                     /db_xref="taxon:562"
                     /plasmid="pO157"
                     /note="RIMD 0509952 is a strain of enterohemorrhagic E.
                     coli, EHEC O157:H7"
     gene            join(92527. .92721,1. .2502)
                     /gene="tagA"
     CDS             join(92527. .92721,1. .2502)
                     /gene="tagA"
                     /codon_start=1
                     /transl_table=11
                     /product="ToxR-regulated lipoprotein"
                     /protein_id="BAA31757.3"
                     /db_xref="GI:4666293"
                     /translation="MNTKMNERWRTPMKLKYLSCTILAPLAIGVFSATAADNNSAIYF
                     NTSQPINDLQGSLAAEVKFAQSQILPAHPKEGDSQPHLTSLRKSLLLVRPVKADDKTP
                     VQVEARDDNNKILGTLTLYPPSSLPDTIYHLDGVPEGGIDFTPHNGTKKIINTVAEVN
                     KLSDASGSSIHSHLTNNALVEIHTANGRWVRDIYLPQGPDLEGKMVRFVSSAGYSSTV
                     FYGDRKVTLSVGNTLLFKYVNGQWFRSGELENNRITYAQHIWSAELPAHWIVPGLNLV
                     IKQGNLSGRLNDIKIGAPGELLLHTIDIGMLTTPRDRFDFAKDKEAHREYFQTIPVSR
                     MIVNNYAPLHLKEVMLPTGELLTDMDPGNGGWHSGTMRQRIGKELVSHGIDNANYGLN
                     STAGLGENSHPYVVAQLAAHNSRGNYANGIQVHGGSGGGGIVTLDSTLGNEFSHEVGH
                     NYGLGHYVDGFKGSVHRSAENNNSTWGWDGDKKRFIPNFYPSQTNEKSCLNNQCQEPF
                     DGHKFGFDAMAGGSPFSAANRFTMYTPNSSAIIQRFFENKAVFDSRSSTGFSKWNADT
                     QEMEPYEHTIDRAEQITASVNELSESKMAELMAEYAVVKVHMWNGNWTRNIYIPTASA
```

DNRGSILTINHEAGYNSYLFINGDEKVVSQGYKKSFVSDGQFWKERDVVDTREARKPE
QFGVPVTTLVGYYDPEGTLSSYIYPAMYGAYGFTYSDDSQNLSDNDCQLQVDTKEGQL
RFRLANHRANNTVMNKFHINVPTESQPTQATLVCNNKILDTKSLTPAPEGLTYTVNGQ
ALPAKENEGCIVSVNSGKRYCLPVGQRSGYSLPDWIVGQEVYVDSGAKAKVLLSDWDN
LSYNRIGEFVGNVNPADMKKVKAWNGQYLDFSKPRSMRVVYK"

gene 2589. .3464
/gene="etpC"

CDS 2589. .3464
/gene="etpC"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31758.1"
/db_xref="GI:3336999"
/translation="MLFFLSFRGDRGLFIKDIVLKMLTPNRLLCVILLIAGYQLVSVI
HHFWLTQAASVPGLSRVSAPETAVTGDQTEERFVFTLFGRASPLSSEGRAQETMPSLS
DDLLSGEDLDVRGILYSSVAEHSVAIFAHNNRQFSLSVGEKVPSYDATISAIFSDHIV
INYQGKTVSLPLRYDNTEKKNAYDNNNLTVGDVITQDNFRVESVFDIMSFSAVTVNNT
LSGYRLIPGKHSSLFYNAGLHDNDLAVSVNGSELRDTRQAQQIMKQLPELKEIKITVE
RDGQLYDAFIAVGEN"

gene 3675. .5432
/gene="etpD"

CDS 3675. .5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31759.1"
/db_xref="GI:3337000"
/translation="MLNEEQYYQFFLSVLDVYGFAVVDMHNGILKVVRSKDAKTSAVP
VASDVSPGTGDEVVTRVVPVSNVAARDLAPLLRQLNDNAGAGSVVHYEPSNVLLMTGR
AAVMKRLMEIVERVDKVGNRSVATVPLTYASATDVARLVTELTKETDKTAIPAWMTAK
LVADERTNSVLVSGEPISQQRIISIIKQLDRQEDVQGNTKVIYLKYAKAKDLVEVLTG
ISSSIENDSKKSPSTEALRKGVTIKSHEQTNALILTGAPDVIRDLENVISQLDIRRPQ
VLVEAIIAEIQDADGLNLGIQWVNKHAGVAQFTSTGLPITTMVQTRQNEILDSDQSNA
LSMFNGIAAGFYQGNWAMLLTALSTSSKNDILATPSIVTLDNMEATFNVGQEVPVLSG
SQTTSGDNIFNTVERKTVGIKLRVKPQINEGDSVLLEIEQEVSGVADTAVATTTDLGA
TFNTRTVTNAMLVGNGETVVVGGLLDKSIRGSESKVPLLGDIPVLGHLFRAKSEQTAK
RNLMLFIRPTIIRERDGFRHASAEKYQSFNQEQVQSRGKETTALTLNEEQLRLSPDQD
DTAFRKVKAAIAAPYAQEM"

gene 5432. .6937
/gene="etpE"

CDS 5432. .6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"
/translation="MSRVVQNVSESRPLLPFSFSRTQRILLLREQEGNRVFCMEDTPA
SALLEVRRVAEGPLNVTTVSAEAFEKQLVSSYQRDSDEARQMMAEIGNEMDFYTVAGE
LPDREDLLDANDDAPIIRLINAMLTEAIKEKASDIHIETYERHLQVRFRIDGVLREIL
RLHRNLASLLISRIKVMARLDIAEKRVPQDGRMVLRIGGRAVDVRVSTLPSNHGERIV
LRLLDKNSVSLDLAALGMSQQNQRHIDALIRRPHGIILVTGPTGSGKSTTLYAALSLL
NPRDRNIMTVEDPVEYELDGISQTQVNPKVDMTFARSLRAILRQDPDVVLVGEIRDGE
TAQIAVQASLTGHLVLSTLHTNSAAGALSRLQDMGIPPFLLSTSLLAVLAQRLVRTLC
PRCRQPCQVSTELAMDMDIPPETTIWQPAGCQHCSFTGYHGRTGIHELLLIDDRIRTA
IYQGEGELGITRLAGSRYLTLRGDGRQKVLAGETSWEEVVRVTESRLQEEE"

gene 6939. .8162
/gene="etpF"

CDS 6939. .8162
/gene="etpF"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"
/translation="MALFHYQASDIHGRKRSGILEADSARHARQLLREQALIPVRLDE
KQVHHKHSLRSILRFPRPRGGSSAELALLTRQLATLVAASLPLEEALDALLRQSEKPRQ
RNLIAAVRTKVLEGHSLAAAMGMFPGTFERLYCAMVAAGETSGRLDVVLSRLADYTEQ
RQIMRNRLLQALLYPCVLTLVAVGVIAILLTAVVPKVVEQFIHMKQTLPLSTRVLMGA
AEVSQTWGPWLLLAAALGGIAGRMILHQPSQRLAFHHLLLRLPVVGRISRGLNTARYA
RTLSILNASAVPLLQAMHISGDVLSNDWARHQLATAAELVREGVSLHQALEQTSLFPP
MMRHMIASGENSGELDSMLERAADNQDREFSTQMQLALGLFEPLLVVGMAGVVLFIVL
AILQPLLQLNNMMNM"

gene 8193. .8627
/gene="etpG"

CDS 8193. .8627
/gene="etpG"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31762.1"
/db_xref="GI:3337003"
/translation="MRKQHQRGFTLLEIMVVIVILGVLASLVVPNLMGNKDKADRQKV
MSDLVALESTLDMYRLDNNRYPTTEQGLRALVSKPTVQPEPRNYRQDGYIRRLPQDPW
GGDYQLLNPGQYSDIDIFSPGPDGVPNTEDDIGNWTLGNAQP"

gene 8624. .9178
/gene="etpH"

CDS 8624. .9178
/gene="etpH"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337004"
/translation="MSQRGFTILEMMLILLLMGITAGLVLMSFPDSAQNHLQQQRERL
QAQLDYALDRSQQDGLLMGIQVRFDGWKFKVLQRGTAESPPTLAEGGDIWQGYVWQTW
QPRRAAMGGKLPDDVRLELQYLRGLQWMSEHDDGAEPDILLLPGGEVTPFRLLFRQVG
EEAVVGLQVDENGLMTLFEGEVSL"

gene 9193. .9540
/gene="etpI"

CDS 9193. .9540
/gene="etpI"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31764.1"
/db_xref="GI:3337005"
/translation="MTLLEVIVALVVFALAGMALMQASTQQAAGIGRMEEKVLAGWLA
DNQMVQLQLEKTWPENGWGEKTISFAGTEWYLRWGGPDSDVPQPRSLEVEVRRTKEET
SALVSLRSSVVRE"

gene 9537. .10136
/gene="etpJ"

CDS 9537. .10136
/gene="etpJ"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31765.1"
/db_xref="GI:3337006"
/translation="MSQQRVKGFTLLEMLLALAVFAALSISAFQVLQGGIRAHELSRD
KVQRLAELQRGISQMERDLTQMLPRHSRGNEGLLLAAPHLLKSDDWGISFTRNSWLNP
AGMLPRPELQWVGYRLRQQKLERLSYFHVDHPSGVSPDVRVMLDGVHAFRLRFFVNGD
WQARWDSTGILPQAVEVTLVMDDFAELPRLFLVSKETAE"

gene 10133. .11110
/gene="etpK"

CDS 10133. .11110
/gene="etpK"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31766.1"
/db_xref="GI:3337007"
/translation="MKLREQGVALLVVLLILSLMVTIAAVIAERNGRTFLRTVAQLDQ
LQAKWDGYTAETIAKQILQRSRQESPRKTHLAQNWAQSERQFETRGGDVRGQIVDAQA
CFNLNAINYGVVDLTSIPYAARIFQQLLINLQVELLQARQVTAALRDWIDRDDKPVRG
GAEDEVYMGMEPPYLAANQPMQDVSELRLIRGIDARLYRKLLPYVCVLPTSDLSVNVN
TLLDSQAPLLAALFLTKPDSLPVTELLQRRPRTGWESVAAFLDAPPLKDIDTSAAMPV
LAVSSNYFLVRLHVRSGEHLFSQQTLMQWREERFRIIQRQYGLTMREVP"

gene 11473. .12321

Query Match 99.2%; Score 1476.4; DB 1; Length 92721;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy        1 CTGCAGTCCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
Db    76351 CTGCAGTCCGGAGATGAACAGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 76410

Qy       60 ATGATTTTTGTCACAGACTCAATGACTACCGGACGCACTGAACCTTCCGGTTGTTTCTCC 119
Db    76411 ATGATTTTTGTCACAGACTCAATGACTACCGGACGCACTGAACCTTCCGGTTGTTTCTCC 76470

Qy      120 AGCCAGTTAAGCCAGCGGTTTCCCTGCTGAAAAATGTCGGCAAAACGGGGAAGCATCAGA 179
Db    76471 AGCCAGTTAAGCCAGCGGTTTCCCTGCTGAAAAATGTCGGCAAAACGGGGAAGCATCAGA 76530

Qy      180 AGGGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC 239
Db    76531 AGGGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC 76590

Qy      240 CGGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAGCGGGCTATTTCAG 299
Db    76591 CGGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAGCGGGCTATTTCAG 76650

Qy      300 GATACCCTTCGTCATCAACACGTACAAACCAGAAGACCAGCTTTTTGTTTCTGACATCCA 359
Db    76651 GATACCCTTCGTCATCAACACGTACAAACCAGAAGACCAGCTTTTTGTTTCTGACATCCA 76710

Qy      360 CAAAGAAGGGAATATTCAGGTCTGCGCAGCACTCAACGGCATCGTCAGTTGCGGCTTGGA 419
Db    76711 CAAAGAAGGGAATATTCAGGTCTGCGCAGCACTCAACGGCATCGTCAGTTGCGGCTTGGA 76770

Qy      420 ACCCCTTAGTATTTTTTGTCTGTAGTATCTATCCCAGCAATAGGTATATCCTGTTGCATC 479
Db    76771 ACCCCTTAGTATTTTTTGTCTGTAGTATCTATCCCAGCAATAGGTATATCCTGTTGCATC 76830

Qy      480 AATAAAGTTGACTTTTGTATACAACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATG 539
Db    76831 AATAAAGTTGACTTTTGTATACAACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATG 76890

Qy      540 ATAAAAAAAACTCTTCCTGTTCTGATTCTTCTGGCGCTATCGGGGAGCTTTTCTACCGCT 599
Db    76891 ATAAAAAAAACTCTTCCTGTTCTGATTCTTCTGGCGCTATCGGGGAGCTTTTCTACCGCT 76950

Qy      600 GTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT 659
Db    76951 GTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT 77010

Qy      660 CCTCTGAGATTACACAGCCCTGAATCAAATCCCTGGGGGGCTGATTTTGATTATGCCACC 719
Db    77011 CCTCTGAGATTACACAGCCCTGAATCAAATCCCTGGGGGGCTGATTTTGATTATGCCACC 77070

Qy      720 AGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAAGATTTGCTGACAACT 779
Db    77071 AGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAAGATTTGCTGACAACT 77130

Qy      780 TCCCAGGATTGGTGCCCTGCGGATTATGGTCATTATGGTCCTTTCTTTATTCGTATGGCT 839
Db    77131 TCCCAGGATTGGTGGCCTGCGGATTATGGTCATTATGGTCCTTTCTTTATTCGTATGGCT 77190

Qy      840 TGGCACGGTGCCGGAACATACAGGACATATGATGGCCGGGGAGGCGCCAGTGGTGGTCAG 899
Db    77191 TGGCACGGTGCCGGAACATACAGGACATATGATGGCCGGGGAGGCGCCAGTGGTGGTCAG 77250

Qy      900 CAACGTTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCGA 959
Db    77251 CAACGTTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCGA 77310

Qy      960 TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCCTGGGGAGACCTGATGGTC 1019
Db    77311 TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCCTGGGGAGACCTGATGGTC 77370

Qy     1020 CTGACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAAACGCTGGGATTTGCTGGCGGA 1079
Db    77371 CTGACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAAACGCTGGGATTTGCTGGCGGA 77430

Qy     1080 AGAGAAGATGACTGGGAGTCGGACCTGGTATACTGGGGGCCTGACAACAAGCCTCTTGCA 1139
Db    77431 AGAGAAGATGACTGGGAGTCGGACCTGGTATACTGGGGGCCTGACAACAAGCCTCTTGCA 77490

Qy     1140 GATAACCGGGATAAAAACGGGAAACTTCAGAAACCTCTTGCCGCCACGCAGATGGGACTT 1199
Db    77491 GATAACCGGGATAAAAACGGGAAACTTCAGAAACCTCTTGCCGCCACGCAGATGGGACTT 77550

Qy     1200 ATTTATGTCAATCCTGAAGGCCCCGGTGGAAAACCAGATCCTCTGGCTTCCGCGAAAGAT 1259
Db    77551 ATTTATGTCAATCCTGAAGGCCCCGGTGGAAAACCAGATCCTCTGGCTTCCGCGAAAGAT 77610

Qy     1260 ATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG 1319
Db    77611 ATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG 77670

Qy     1320 GGAGGGCATACATTTGGTAAAGCACATGGTGCAGCGTCTCCTGAAAAATGTATTGGCGCA 1379
Db    77671 GGAGGGCATACATTTGGTAAAGCACATGGTGCAGCGTCTCCTGAAAAATGTATTGGCGCA 77730

Qy     1380 GGGCCTGATGGTGCACCTGTGGAGGAGCAGGGACTGGGATGGAAAAATAAATGTGGTACA 1439
Db    77731 GGGCCTGATGGTGCACCTGTGGAGGAGCAGGGACTGGGATGGAAAAATAAATGTGGTACA 77790

Qy     1440 GGAAACGGCAAATATACCATCACCAGTGGCCTGGAAGGAGCCTGGTCGAC 1489
Db    77791 GGAAACGGCAAATATACCATCACCAGTGGCCTGGAAGGAGCCTGGTCGAC 77840
```

```
RESULT 6
AX191725
LOCUS       AX191725               92721 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 7 from Patent WO0149775.
ACCESSION   AX191725
VERSION     AX191725.1  GI:15209894
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 7 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             99.2%;  Score 1476.4;  DB 6;  Length 92721;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1488;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CTGCAGTCCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
Db      76351 CTGCAGTCCGGAGATGAACAGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 76410

Qy         60 ATGATTTTTGTCACAGACTCAATGACTACCGGACGCACTGAACCTTCCGGTTGTTTCTCC 119
Db      76411 ATGATTTTTGTCACAGACTCAATGACTACCGGACGCACTGAACCTTCCGGTTGTTTCTCC 76470

Qy        120 AGCCAGTTAAGCCAGCGGTTTCCCTGCTGAAAAATGTCGGCAAAACGGGGAAGCATCAGA 179
Db      76471 AGCCAGTTAAGCCAGCGGTTTCCCTGCTGAAAAATGTCGGCAAAACGGGGAAGCATCAGA 76530

Qy        180 AGGGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC 239
Db      76531 AGGGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC 76590

Qy        240 CGGCGCTGATACCGGCAAGAATGGTCGCAAACTCCCGCTCCGTGCAGCGGGCTATTTCAG 299
```

```
Qy        971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGTCATGGCCACT 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18464 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGTCATGGCCACT 18405

Qy       1031 ATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGTCTGAGGATGGA 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18404 ATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGTCTGAGGATGGA 18345

Qy       1091 AGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTTCTGGC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18344 AGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTTCTGGC 18285

Qy       1151 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
              ||||||||||||||||||||||||||||||
Db      18284 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 18254
```

RESULT 7
AB011549/c
LOCUS AB011549 92721 bp DNA circular BCT 27-APR-1999
DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION AB011549
VERSION AB011549.2 GI:4589740
KEYWORDS ToxR-regulated lipoprotein; tagA.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (sites)
AUTHORS Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K., Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K., Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S. and Shinagawa,H.
TITLE Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic Escherichia coli O157:H7 derived from Sakai outbreak
JOURNAL DNA Res. 5 (1), 1-9 (1998)
MEDLINE 98290540
PUBMED 9628576
REFERENCE 2 (bases 1 to 92721)
AUTHORS Makino,K.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1998) Kozo Makino, Research Institute for Microbial Diseases, Osaka University, Molecular Microbiology; yamadaoka, 3-1, Suita, Osaka 562, Japan (E-mail:makino@bkns01.biken.osaka-u.ac.jp, Tel:81-6-879-8318, Fax:81-6-879-8320)
COMMENT On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES Location/Qualifiers
source 1. .92721
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:562"
/plasmid="pO157"
/note="RIMD 0509952 is a strain of enterohemorrhagic E. coli, EHEC O157:H7"
gene join(92527. .92721,1. .2502)
/gene="tagA"
CDS join(92527. .92721,1. .2502)
/gene="tagA"
/codon_start=1
/transl_table=11
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4666293"
/translation="MNTKMNERWRTPMKLKYLSCTILAPLAIGVFSATAADNNSAIYF
NTSQPINDLQGSLAAEVKFAQSQILPAHPKEGDSQPHLTSLRKSLLLVRPVKADDKTP
VQVEARDDNNKILGTLTLYPPSSLPDTIYHLDGVPEGGIDFTPHNGTKKIINTVAEVN
KLSDASGSSIHSHLTNNALVEIHTANGRWVRDIYLPQGPDLEGKMVRFVSSAGYSSTV
FYGDRKVTLSVGNTLLFKYVNGQWFRSGELENNRITYAQHIWSAELPAHWIVPGLNLV
IKQGNLSGRLNDIKIGAPGELLLHTIDIGMLTTPRDRPDPAKDKEAHREYFQTIPVSR
MIVNNYAPLHLKEVMLPTGELLTDMDPGNGGWHSGTMRQRIGKELVSHGIDNANYGLN
STAGLGENSHPYVVAQLAAHNSRGNYANGIQVHGGSGGGGIVTLDSTLGNEFSHEVGH
NYGLGHYVDGFKGSVHRSAENNNSTWGWDGDKKRFIPNFYPSQTNEKSCLNNQCQEPF
DGHKFGFDAMAGGSPFSAANRFTMYTPNSSAIIQRFFENKAVFDSRSSTGFSKWNADT
QEMEPYEHTIDRAEQITASVNELSESKMAELMAEYAVVKVHMWNGNWTRNIYIPTASA
DNRGSILTINHEAGYNSYLFINGDEKVVSQGYKKSFVSDGQFWKERDVVDTREARKPE
QFGVPVTTLVGYYDPEGTLSSYIYPAMYGAYGFTYSDDSQNLSDNDCQLQVDTKEGQL
RFRLANHRANNTVMNKFHINVPTESQPTQATLVCNNKILDTKSLTPAPEGLTYTVNGQ
ALPAKENEGCIVSVNSGKRYCLPVGQRSGYSLPDWIVGQEVYVDSGAKAKVLLSDWDN
LSYNRIGEFVGNVNPADMKKVKAWNGQYLDFSKPRSMRVVYK"
gene 2589. .3464
/gene="etpC"
CDS 2589. .3464
/gene="etpC"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related pr
/protein_id="BAA31758.1"
/db_xref="GI:3336999"
/translation="MLFFLSFRGDRGLFIKDIVLKMLTPNRLLCVI
HHFWLTQAASVPGLSRVSAPETAVTGDQTEERFVPTLFGRASPLSS
DDLLSGEDLDVRGILYSSVAEHSVAIFAHNNRQFSLSVGEKVPSYD
INYQGKTVSLPLRYDNTEKKNAYDNNNLTVGDVITQDNFRVESVFD
LSGYRLIPGKHSSLFYNAGLHDNDLAVSVNGSELRDTRQAQQIMKQ
RDGQLYDAFIAVGEN"
gene 3675. .5432
/gene="etpD"
CDS 3675. .5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related pro
/protein_id="BAA31759.1"
/db_xref="GI:3337000"
/translation="MLNEEQYYQFFLSVLDVYGFAVVDMHNGILKVVRSKDAKTSAVP
VASDVSPGTGDEVVTRVVPVSNVAARDLAPLLRQLNDNAGAGSVVHYEPSNVLLMTGR
AAVMKRLMEIVERVDKVGNRSVATVPLTYASATDVARLVTELTKETDKTAIPAWMTAK
LVADERTNSVLVSGEPISQQRIISIIKQLDRQEDVQGNTKVIYLKYAKAKDLVEVLTG
ISSSIENDSKKSPSTEALRKGVTIKSHEQTNALILTGAPDVIRDLENVISQLDIRRPQ
VLVEAIIAEIQDADGLNLGIQWVNKHAGVAQFTSTGLPITTMVQTRQNEILDSDQSNA
LSMFNGIAAGFYQGNWAMLLTALSTSSKNDILATPSIVTLDNMEATFNVGQEVPVLSG
SQTTSGDNIFNTVERKTVGIKLRVKPQINEGDSVLLEIEQEVSGVADTAVATTTDLGA
TFNTRTVTNAMLVGNGETVVVGGLLDKSIRGSESKVPLLGDIPVLGHLFRAKSEQTAK
RNLMLFIRPTIIRERDGFRHASAEKYQSFNQEQVQSRGKETTALTLNEEQLRLSPDQD
DTAFRKVKAAIAAFYAQEM"
gene 5432. .6937
/gene="etpE"
CDS 5432. .6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"
/translation="MSRVVQNVSESRPLLPFSFSRTQRILLLREQEGNRVFCMEDTPA
SALLEVRRVAEGPLNVTTVSAEAFEKQLVSSYQRDSDEARQMMAEIGNEMDFYTVAGE
LPDREDLLDANDDAPIIRLINAMLTEAIKEKASDIHIETYERHLQVRFRIDGVLREIL
RLHRNLASLLISRIKVMARLDIAEKRVPQDGRMVLRIGGRAVDVRVSTLPSNHGERIV
LRLLDKNSVSLDLAALGMSQQNQRHIDALIRRPHGIILVTGPTGSGKSTTLYAALSLL
NPRDRNIMTVEDPVEYELDGISQTQVNPKVDMTFARSLRAILRQDPDVVLVGEIRDGE
TAQIAVQASLTGHLVLSTLHTNSAAGALSRLQDMGIPPFLLSTSLLAVLAQRLVRTLC
PRCRQPCQVSTELAMDMDIPPETTIWQPAGCQHCSFTGYHGRTGIHELLLIDDRIRTA
IYQGEGELGITRLAGSRYLTLRGDGRQKVLAGETSWEEVVRVTESRLQEEE"
gene 6939. .8162
/gene="etpF"
CDS 6939. .8162
/gene="etpF"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"

```
                /translation="MALFHYQASDIHGRKRSGILEADSARHARQLLREQALIPVRLDE
                KQVHHKHSLRSILRFRPRGGSSAELALLTRQLATLVAASLPLEEALDALLRQSEKPRQ
                RNLIAAVRTKVLEGHSLAAAMGMFPGTFERLYCAMVAAGETSGRLDVVLSRLADYTEQ
                RQIMRNRLLQALLYPCVLTLVAVGVIAILLTAVVPKVVEQFIHMKQTLPLSTRVLMGA
                AEVSQTWGPWLLLAAALGGIAGRMILHQPSQRLAFHHLLLRLPVVGRISRGLNTARYA
                RTLSILNASAVPLLQAMHISGDVLSNDWARHQLATAAELVREGVSLHQALEQTSLFPP
                MMRHMIASGENSGELDSMLERAADNQDREFSTQMQLALGLFEPLLVVGMAGVVLFIVL
                AILQPLLQLNNMMNM"
gene            8193. .8627
                /gene="etpG"
CDS             8193. .8627
                /gene="etpG"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31762.1"
                /db_xref="GI:3337003"
                /translation="MRKQHQRGFTLLEIMVVIVILGVLASLVVPNLMGNKDKADRQKV
                MSDLVALESTLDMYRLDNNRYPTTEQGLRALVSKPTVQPEPRNYRQDGYIRRLPQDPW
                GGDYQLLNPGQYSDIDIFSPGPDGVPNTEDDIGNWTLGNAQP"
gene            8624. .9178
                /gene="etpH"
CDS             8624. .9178
                /gene="etpH"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31763.1"
                /db_xref="GI:3337004"
                /translation="MSQRGFTILEMMLILLLMGITAGLVLMSFPDSAQNHLQQQRERL
                QAQLDYALDRSQQDGLLMGIQVRFDGWKFKVLQRGTAESPPTLAEGGDIWQGYVWQTW
                QPRRAAMGGKLPDDVRLELQYLRGLQWMSEHDDGAEPDILLLPGGEVTPFRLLFRQVG
                EEAVVGLQVDENGLMTLFEGEVSL"
gene            9193. .9540
                /gene="etpI"
CDS             9193. .9540
                /gene="etpI"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31764.1"
                /db_xref="GI:3337005"
                /translation="MTLLEVIVALVVFALAGMALMQASTQQAAGIGRMEEKVLAGWLA
                DNQMVQLQLEKTWPENGWGEKTISFAGTEWYLRWGGPDSDVPQPRSLEVEVRRTKEET
                SALVSLRSSVVRE"
gene            9537. .10136
                /gene="etpJ"
CDS             9537. .10136
                /gene="etpJ"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31765.1"
                /db_xref="GI:3337006"
                /translation="MSQQRVKGFTLLEMLLALAVFAALSISAFQVLQGGIRAHELSRD
                KVQRLAELQRGISQMERDLTQMLPRHSRGNEGLLLAAPHLLKSDDWGISFTRNSWLNP
                AGMLPRPELQWVGYRLRQQKLERLSYFHVDHPSGVSPDVRVMLDGVHAFRLRFFVNGD
                WQARWDSTGILPQAVEVTLVMDDFAELPRLFLVSKETAE"
gene            10133. .11110
                /gene="etpK"
CDS             10133. .11110
                /gene="etpK"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31766.1"
                /db_xref="GI:3337007"
                /translation="MKLREQGVALLVVLLILSLMVTIAAVIAERNGRTFLRTVAQLDQ
                LQAKWDGYTAETIAKQILQRSRQESPRKTHLAQNWAQSERQFETRGGDVRGQIVDAQA
                CFNLNAINYGVVDLTSIPYAARIFQQLLINLQVELLQARQVTAALRDWIDRDDKPVRG
                GAEDEVYMGMEPPYLAANQPMQDVSELRLIRGIDARLYRKLLPYVCVLPTSDLSVNVN
                TLLDSQAPLLAALFLTKPDSLPVTELLQRRPRTGWESVAAFLDAPPLKDIDTSAAMPV
                LAVSSNYFLVRLHVRSGEHLFSQQTLMQWREERFRIIQRQYGLTMREVP"
```

```
gene            11473. .12321

 Query Match           98.7%;  Score 1166.2;  DB 1;  Length 92721;
 Best Local Similarity  99.7%;  Pred. No. 0;
 Matches 1168;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        11 TGGAAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88939 TGGAAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCTGAAAGTAGGCC 88880

Qy        71 TGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAATGTGATAAATTA 130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88879 TGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAATGTGATAAATTA 88820

Qy       131 CTCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCCGTTATCT 190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88819 CTCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCCGTTATCT 88760

Qy       191 GTTGCATTTCCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTTAAT 250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88759 GTTGCATTTCCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTTAAT 88700

Qy       251 ATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88699 ATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 88640

Qy       311 TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCCACGGGTAACAG 370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88639 TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCCACGGGTAACAG 88580

Qy       371 CGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAAC 430
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88579 CGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAAC 88520

Qy       431 GGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAACGCTTCATTTTCCGAAGG 490
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88519 GGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAACGCTTCATTTTCCGAAGG 88460

Qy       491 CATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGGGCACCTGCCAT 550
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88459 CATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGGGCACCTGCCAT 88400

Qy       551 TAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGTGCTTCGTATTC 610
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88399 TAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGTGCTTCGTATTC 88340

Qy       611 TTCACGCCCGGCGCCCACCACGAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA 670
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88339 TTCACGCCCGGCGCCCACCACGAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA 88280

Qy       671 CAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88279 CAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 88220

Qy       731 AAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGTAACGGTCGCAT 790
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88219 AAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGTAACGGTCGCAT 88160

Qy       791 GGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAATAACCGTTACCGGTGTCCT 850
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88159 GGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAATAACCGTTACCGGTGTCCT 88100

Qy       851 GACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACAATGAAATCACA 910
           |||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||
Db     88099 GACACCTTCCGCCATCAGATGCGCCATCATGGCTGAGCTAGGCACAACAATGAAATCACA 88040

Qy       911 CAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTTTGTCTGACAAT 970
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88039 CAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTTTGTCTGACAAT 87980

Qy       971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGTCATGGCCACT 1030
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87979 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGTCATGGCCACT 87920
```

```
Qy       1031 ATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGTCTGAGGATGGA 1090
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87919 ATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGTCTGAGGATGGA 87860

Qy       1091 AGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTTCTGGC 1150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87859 AGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTTCTGGC 87800

Qy       1151 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
             ||||||||||||||||||||||||||||||
Db      87799 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 87769


RESULT 8
AX191725/c
LOCUS       AX191725               92721 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 7 from Patent WO0149775.
ACCESSION   AX191725
VERSION     AX191725.1  GI:15209894
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 7 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             98.7%;  Score 1166.2;  DB 6;  Length 92721;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1168;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         11 TGGAAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88939 TGGAAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCTGAAAGTAGGCC 88880

Qy         71 TGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAATGTGATAAATTA 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88879 TGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAATGTGATAAATTA 88820

Qy        131 CTCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCCGTTATCT 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88819 CTCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCCGTTATCT 88760

Qy        191 GTTGCATTTCCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTTAAT 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88759 GTTGCATTTCCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTTAAT 88700

Qy        251 ATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88699 ATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 88640

Qy        311 TCGGGATAAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCCACGGGTAACAG 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88639 TCGGGATAAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCCACGGGTAACAG 88580

Qy        371 CGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAAC 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88579 CGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAAC 88520

Qy        431 GGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAACGCTTCATTTTCCGAAGG 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88519 GGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAACGCTTCATTTTCCGAAGG 88460
```

```
Qy         491 CATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGGGCACCTGCCAT 550
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88459 CATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGGGCACCTGCCAT 88400

Qy         551 TAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGTGCTTCGTATTC 610
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88399 TAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGTGCTTCGTATTC 88340

Qy         611 TTCACGCCCGGCGCCCACCACGAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA 670
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88339 TTCACGCCCGGCGCCCACCACGAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA 88280

Qy         671 CAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88279 CAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 88220

Qy         731 AAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGTAACGGTCGCAT 790
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88219 AAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGTAACGGTCGCAT 88160

Qy         791 GGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAATAACCGTTACCGGTGTCCT 850
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88159 GGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAATAACCGTTACCGGTGTCCT 88100

Qy         851 GACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACAATGAAATCACA 910
               ||||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db       88099 GACACCTTCCGCCATCAGATGCGCCATCATGGCTGAGCTAGGCACAACAATGAAATCACA 88040

Qy         911 CAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTTTGTCTGACAAT 970
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       88039 CAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTTTGTCTGACAAT 87980

Qy         971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGTCATGGCCACT 1030
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       87979 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGTCATGGCCACT 87920

Qy        1031 ATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGTCTGAGGATGGA 1090
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       87919 ATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGTCTGAGGATGGA 87860

Qy        1091 AGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTTCTGGC 1150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       87859 AGGAAGGTGAAGGCTGTTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTTCTGGC 87800

Qy        1151 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
               ||||||||||||||||||||||||||||||
Db       87799 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 87769


RESULT 9
AF043470
LOCUS       AF043470                5612 bp    DNA     linear   BCT 25-JUN-1998
DEFINITION  Escherichia coli plasmid pO157 ecf4 gene, partial cds; and ecf3,
            ecf2, and ecf1 genes, complete cds.
ACCESSION   AF043470
VERSION     AF043470.1  GI:3253288
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 5612)
  AUTHORS   Boerlin,P., Chen,S., Colbourne,J.K., Johnson,R., De Grandis,S. and
            Gyles,C.
  TITLE     Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids
            and the locus for enterocyte effacement in shiga toxin-producing E.
            coli
  JOURNAL   Infect. Immun. 66 (6), 2553-2561 (1998)
  MEDLINE   98261495
   PUBMED   9596716
REFERENCE   2  (bases 1 to 5612)
  AUTHORS   Boerlin,P. and Gyles,C.
  TITLE     Direct Submission
```

```
     source          1. .14
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db         11 ATCGTCAG 4


RESULT 7
AR126298/c
LOCUS       AR126298                  14 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 5 from patent US 6180339.
ACCESSION   AR126298
VERSION     AR126298.1  GI:14112891
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 14)
  AUTHORS   Sandhu,G.S. and Kline,B.C.
  TITLE     Nucleic acid probes for the detection and identification of fungi
  JOURNAL   Patent: US 6180339-A 5 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .14
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db         11 ATCGTCAG 4


RESULT 8
BD132864/c
LOCUS       BD132864                  14 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Nucleic acid probes for the detection and identification of fungi.
ACCESSION   BD132864
VERSION     BD132864.1  GI:23227809
KEYWORDS    JP 2002504817-A/5.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 14)
  AUTHORS   Sandhu,G.S. and Kline,B.C.
  TITLE     Nucleic acid probes for the detection and identification of fungi
  JOURNAL   Patent: JP 2002504817-A 5 12-FEB-2002;
            BAYER CORP
COMMENT     PN   JP 2002504817-A/5
            PD   12-FEB-2002
            PF   04-JUN-1998 JP 1999501953
            PR   06-JUN-1997 US    08/871678
            PI   GURPREET S SANDHU,BRUCE C KLINE
            PC   C12Q1/68
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .14
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db         11 ATCGTCAG 4


RESULT 9
BD205228
LOCUS       BD205228                  14 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleotide sequence for detecting enterohemorrhagic Escherichia
            coli (EHEC).
ACCESSION   BD205228
VERSION     BD205228.1  GI:33014998
KEYWORDS    JP 2002512813-A/18.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 14)
  AUTHORS   Frechon,D.T.M., Laure,F.C. and Thierry,D.
  TITLE     Nucleotide sequence for detecting enterohemorrhagic Escherichia
            coli (EHEC)
  JOURNAL   Patent: JP 2002512813-A 18 08-MAY-2002;
            BIORAD PASTEUR
COMMENT     OS   Unidentified
            PN   JP 2002512813-A/18
            PD   08-MAY-2002
            PF   27-APR-1999 JP 2000546051
            PR   28-APR-1998 FR     98/05329
            PI   DOMINIQUE THERESE MARIE FRECHON,FRANCOISE CLAUDINE LAURE, PI
            DOMINIQUE THIERRY
            PC   C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Nucleotide sequence for detecting enterohemorrhagic  CC
                Escherichia coli
            CC   (EHEC).
            FH   Key             Location/Qualifiers
            FT   source          1. .14
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .14
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db          4 ATCGTCAG 11


RESULT 10
I79345/c
LOCUS       I79345                    14 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 5 from patent US 5707802.
ACCESSION   I79345
VERSION     I79345.1  GI:3207635
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 14)
  AUTHORS   Sandhu,G.S. and Kline,B.C.
  TITLE     Nucleic acid probes for the detection and identification of fungi
  JOURNAL   Patent: US 5707802-A 5 13-JAN-1998;
```

```
FEATURES             Location/Qualifiers
     source          1. .14
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db         11 ATCGTCAG 4


RESULT 11
AX011314
LOCUS       AX011314                  14 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 18 from Patent WO9955908.
ACCESSION   AX011314
VERSION     AX011314.1  GI:9997864
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Thierry,D., Frechon,D.T. and Laure,F.C.
  TITLE     Nucleotide sequences for detecting enterohemorrhagic escherichia
            coli (ehec)
  JOURNAL   Patent: WO 9955908-A 18 04-NOV-1999;
            THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
            FRANCOISE CLAUDINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES             Location/Qualifiers
     source          1. .14
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db          4 ATCGTCAG 11


RESULT 12
AR301299/c
LOCUS       AR301299                  15 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 61 from patent US 6538122.
ACCESSION   AR301299
VERSION     AR301299.1  GI:31689072
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 15)
  AUTHORS   Yokoyama,K., Nakamura,N., Miwa,T. and Seguro,K.
  TITLE     Process for producing microbial transglutaminase
  JOURNAL   Patent: US 6538122-A 61 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .15
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db          9 ATCGTCAG 2


RESULT 13
AX103974/c
LOCUS       AX103974                  15 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 166 from Patent WO0122972.
ACCESSION   AX103974
VERSION     AX103974.1  GI:13920171
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 166 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .15
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db         10 ATCGTCAG 3


RESULT 14
AX355608/c
LOCUS       AX355608                  15 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 636 from Patent WO0197843.
ACCESSION   AX355608
VERSION     AX355608.1  GI:18620276
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Weiner,G. and Hartmann,G.
  TITLE     Methods for enhancing antibody-induced cell lysis and treating
            cancer
  JOURNAL   Patent: WO 0197843-A 636 27-DEC-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .15
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide-phosphodiester backbone"
ORIGIN

  Query Match             100.0%;  Score 8;  DB 6;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCGTCAG 8
              ||||||||
Db         10 ATCGTCAG 3


RESULT 15
AX547027/c
LOCUS       AX547027                  15 bp    DNA     linear   PAT 01-MAR-2003
```

nd Specific

```
                     LAHYTNGATLRFTYLDHRTQAYQQETLSQADMLFRVVQHIPEKHFRMIRYFGFLANRV
                     CGQYLPKVYEALKMATPGPTPKLYFAPMAKAFLNVDPFRCVLCGARMVYTAAISGLTV
                     QGLNLNAQAIAQMRYVKP"
     RBS             387. .391
                     /note="putative"
     CDS             complement(461. .1060)
                     /note="unnamed protein product; orf2"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA40541.1"
                     /db_xref="GI:45832"
                     /db_xref="UniProt/TrEMBL:Q99178"
                     /translation="MLFRDVLHHPEKHVGLAQRFLLIGLCAVIQVGEPQRGPVGVMRQ
                     TTARDGRFFQVAAQVIDGLSAVFRLLRQVDVPVLATSGQYQTAPVAFTEDVGQRRRHR
                     QLRKRLSQQVIAHQPHPAPPQGLFVERQIFPDALLIQAARGHRQVDVGMPVEPPAIGV
                     QRPENPDAQTAPLGGVDQVIGRQPKQRIEQPAVVEEQGP"
     misc_feature    1513. .1517
                     /note="insertion target sequence"
                     /evidence=experimental
ORIGIN

  Query Match             100.0%;  Score 14;  DB 1;  Length 1517;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGCATCGTCAGTTG 14
              ||||||||||||||
Db        762 GGCATCGTCAGTTG 749


RESULT 19
AB022426
LOCUS       AB022426               2055 bp    mRNA    linear   MAM 21-JAN-1999
DEFINITION  Sus scrofa mRNA for FXII, complete cds.
ACCESSION   AB022426
VERSION     AB022426.1  GI:4165316
KEYWORDS    FXII.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2055)
  AUTHORS   Takahashi,T. and Kihara,T.
  TITLE     Porcine liver factor XII
  JOURNAL   Published Only in DataBase (1999)
REFERENCE   2  (bases 1 to 2055)
  AUTHORS   Takahashi,T. and Kihara,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-1999) Takayuki Takahashi, Hokkaido University,
            Graduate School of Science; Kitaku Kita 10 joh Nishi 8 choume,
            Sapporo, Hokkaido 060-0810, Japan
            (E-mail:ttakaha@sci.hokudai.ac.jp, Tel:81-11-706-2748,
            Fax:81-11-706-2748)
FEATURES             Location/Qualifiers
     source          1. .2055
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="liver"
     CDS             25. .1875
                     /codon_start=1
                     /product="FXII"
                     /protein_id="BAA37148.1"
                     /db_xref="GI:4165317"
                     /translation="MRALLLLGILLVSLESALLIPPWKDPRKHKVMASEHTVVLTVTG
                     EPCHFPFQYYRQLYYKCIQRGQRGPRPWCATTPNFEKDQRWAYCLEPMKVKDHCNKGN
                     PCQKGGTCVNMPNGPHCICPDHFTGKHCQKEKCFEPQFLQFFQENEIWHRFEPAGVSK
                     CQCKGPKAQCKPVASQVCSTNPCLNGGSCLQTEGHRLCRCPTGYAGRLCDVDLKERCY
                     SDRGLSYRGMAQTTLSGAPCQPWASEATYWNMTAEQALNWGLGDHAFCRNPDNDTRPW
                     CFVWRGDQLSWQYCRLARCQAPIGEAPPILTPTQSPSEHQDSPLLSREPQPTTQTPSQ
                     NLTSAWCAPPEQRGPLPSAGLVGCGQRLRKRLSSLNRIVGGLVALPGAHPYIAALYWG
                     QNFCAGSLIAPCWVLTAAHCLQNRPAPEELTVVLGQDRHNQSCEQCQTLAVRSYRLHE
                     SYSPKTYQHDLALVRLKETADGCCAHPSPFVQPVCLPRSVASSAEPEGALCEVAGWGH
                     QFEGAEEYSSFLQEAQVPLISPERCSAADVHGAAFTPGMLCAGFLEGGTDACQGDSGG
                     PLVCEDETAERQLVLRGIVSWGSGCGDRLKPGVYTDVANYLAWIQEHTTS"
ORIGIN

  Query Match             100.0%;  Score 14;  DB 4;  Length 2055;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps

Qy          1 GGCATCGTCAGTTG 14
              ||||||||||||||
Db       1771 GGCATCGTCAGTTG 1784


RESULT 20
AX073964/c
LOCUS       AX073964               5145 bp    DNA     linear   PAT 06-FEB-20
DEFINITION  Sequence 1 from Patent WO0104295.
ACCESSION   AX073964
VERSION     AX073964.1  GI:12710224
KEYWORDS    .
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Avides,M.D., Deak,P. and Glover,D.M.
  TITLE     Orbit and homologues thereof
  JOURNAL   Patent: WO 0104295-A 1 18-JAN-2001;
            University of Dundee (GB)
FEATURES             Location/Qualifiers
     source          1. .5145
                     /organism="Drosophila melanogaster"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:7227"
ORIGIN

  Query Match             100.0%;  Score 14;  DB 6;  Length 5145;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGCATCGTCAGTTG 14
              ||||||||||||||
Db       3749 GGCATCGTCAGTTG 3736


RESULT 21
AF195498/c
LOCUS       AF195498               5916 bp    mRNA    linear   INV 01-NOV-2000
DEFINITION  Drosophila melanogaster clone LD11488 Misexpression suppressor of
            ras 7 (MESR7) mRNA, MESR7-3403 allele, complete cds.
ACCESSION   AF195498
VERSION     AF195498.1  GI:11066120
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 5916)
  AUTHORS   Huang,A.M., Rubin,G.M., Tsang,G., Evans-Holm,M. and Suh,C.
  TITLE     Full length Drosophila melanogaster cDNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5916)
  AUTHORS   Huang,A.M., Rubin,G.M., Tsang,G., Evans-Holm,M. and Suh,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-1999) Molecular and Cell Biology, University of
            California at Berkeley, 545 Life Sciences Addition Bldg., Berkeley,
            CA 94720-3200, USA
FEATURES             Location/Qualifiers
     source          1. .5916
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
```

specificity

Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCATCGTCAGTTG 14
              ||||||||||||||
Db        420 GGCATCGTCAGTTG 407
```

RESULT 2
US-10-282-122A-36930
; Sequence 36930, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36930
; LENGTH: 1285
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-36930

Query Match 100.0%; Score 14; DB 17; Length 1285;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCATCGTCAGTTG 14
              ||||||||||||||
Db        403 GGCATCGTCAGTTG 416
```

RESULT 3
US-10-282-122A-40133
; Sequence 40133, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40133
; LENGTH: 1551
; TYPE: DNA
; ORGANISM: Salmonella typhi
US-10-282-122A-40133

Query Match 100.0%; Score 14; DB 17; Length 1551;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCATCGTCAGTTG 14
              ||||||||||||||
Db        694 GGCATCGTCAGTTG 707
```

RESULT 4
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484

NOT SPECIFIC

```
LOCUS       BD205211                1489 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleotide sequence for detecting enterohemorrhagic Escherichia
            coli (EHEC).
ACCESSION   BD205211
VERSION     BD205211.1  GI:33014981
KEYWORDS    JP 2002512813-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1489)
  AUTHORS   Frechon,D.T.M., Laure,F.C. and Thierry,D.
  TITLE     Nucleotide sequence for detecting enterohemorrhagic Escherichia
            coli (EHEC)
  JOURNAL   Patent: JP 2002512813-A 1 08-MAY-2002;
            BIORAD PASTEUR
COMMENT     OS   Unidentified
            PN   JP 2002512813-A/1
            PD   08-MAY-2002
            PF   27-APR-1999 JP 2000546051
            PR   28-APR-1998 FR     98/05329
            PI   DOMINIQUE THERESE MARIE FRECHON,FRANCOISE CLAUDINE LAURE, PI
            DOMINIQUE THIERRY
            PC   C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC
            Strandedness: Double;
            CC   Topology: Linear;
            CC   Nucleotide sequence for detecting enterohemorrhagic  CC
               Escherichia coli
            CC   (EHEC).
            FH   Key             Location/Qualifiers
            FT   source          1. .1489
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1489
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 16;  DB 6;  Length 1489;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGCATCGTCAGTTGC 16
              ||||||||||||||||
Db        396 CGGCATCGTCAGTTGC 411


RESULT 10
AX011297
LOCUS       AX011297                1489 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955908.
ACCESSION   AX011297
VERSION     AX011297.1  GI:9997847
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Thierry,D., Frechon,D.T. and Laure,F.C.
  TITLE     Nucleotide sequences for detecting enterohemorrhagic escherichia
            coli (ehec)
  JOURNAL   Patent: WO 9955908-A 1 04-NOV-1999;
            THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
            FRANCOISE CLAUDINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES             Location/Qualifiers
     source          1. .1489
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 16;  DB 6;  Length 1489;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGCATCGTCAGTTGC 16
              ||||||||||||||||
Db        396 CGGCATCGTCAGTTGC 411


RESULT 11
PSIS801/c
LOCUS       PSIS801                 1517 bp    DNA     linear   BCT 07-JUL-2002
DEFINITION  P.syringae DNA for IS801 insertion sequence.
ACCESSION   X57269
VERSION     X57269.1  GI:45830
KEYWORDS    insertion element IS801.
SOURCE      Pseudomonas syringae
  ORGANISM  Pseudomonas syringae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1517)
  AUTHORS   Romantschuk,M., Richter,G.Y., Mukhopadhyay,P. and Mills,D.
  TITLE     IS801, an insertion sequence element isolated from Pseudomonas
            syringae pathovar phaseolicola
  JOURNAL   Mol. Microbiol. 5 (3), 617-622 (1991)
  MEDLINE   91260445
   PUBMED   1646375
REFERENCE   2  (bases 1 to 1517)
  AUTHORS   Mills,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-1991) D. Mills, Oregon State University, Dept of
            Botany and Plant Pathology, Corvallis OR 97331-2902, U S A
FEATURES             Location/Qualifiers
     source          1. .1517
                     /organism="Pseudomonas syringae"
                     /mol_type="genomic DNA"
                     /strain="pathovar phaseolicola, strain LR781"
                     /db_xref="taxon:317"
     misc_feature    1. .5
                     /note="insertion target sequence duplication"
                     /evidence=experimental
     repeat_region   6. .1512
                     /insertion_seq="IS801"
     CDS             101. .1333
                     /note="unnamed protein product; orf1"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA40540.1"
                     /db_xref="GI:45831"
                     /db_xref="GOA:P24607"
                     /db_xref="UniProt/Swiss-Prot:P24607"
                     /translation="MKPAYPPLLLQMSPAYTPRPLKNLFTANQCWAHLLEEGGLRDIE
                     VETVTKMLACGTSILGVKHYTCGNHSCPHVKYLCNTCHCRACPSCGKKATDQWITVQN
                     NRLPDCPWQHLVFTLPDTLWPLFFYNRWLLDALFRLAADNLIYAAKRRGLRVGIFGAL
                     HTYGRRLNWHPHVHLSVTAGGLDEQGVWKNLSFHKEALRRRWMWLVRDYLLGQPLSQL
                     TMPPPLAHILCESDWRRLILAAGGQHWHIHLSKKTKNGRKTVNYLGRYLKKPPISGSR
                     LAHYTNGATLRFTYLDHRTQAYQQETLSQADMLFRVVQHIPEKHFRMIRYFGFLANRV
                     CGQYLPKVYEALKMATPGPTPKLYFAPMAKAFLNVDPFRCVLCGARMVYTAAISGLTV
                     QGLNLNAQAIAQMRYVKP"
     RBS             387. .391
                     /note="putative"
     CDS             complement(461. .1060)
                     /note="unnamed protein product; orf2"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA40541.1"
                     /db_xref="GI:45832"
                     /db_xref="UniProt/TrEMBL:Q99178"
                     /translation="MLFRDVLHHPEKHVGLAQRFLLIGLCAVIQVGEPQRGPVGVMRQ
                     TTARDGRFFQVAAQVIDGLSAVFRLLRQVDVPVLATSGQYQTAPVAFTEDVGQRRRHR
                     QLRKRLSQQVIAHQPHPAPPQGLFVERQIFPDALLIQAARGHRQVDVGMPVEPPAIGV
                     QRPENPDAQTAPLGGVDQVIGRQPKQRIEQPAVVEEQGP"
     misc_feature    1513. .1517
                     /note="insertion target sequence"
```

```
                     /evidence=experimental
ORIGIN

  Query Match             100.0%;  Score 16;  DB 1;  Length 1517;
  Best Local Similarity   100.0%;  Pred. No. 6.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGCATCGTCAGTTGC 16
              ||||||||||||||||
Db        763 CGGCATCGTCAGTTGC 748


RESULT 12
AY603980
LOCUS       AY603980               40110 bp    DNA     circular BCT 20-JUL-2004
DEFINITION  Pseudomonas syringae pv. maculicola strain ES4326 plasmid
            pPMA4326B, complete sequence.
ACCESSION   AY603980
VERSION     AY603980.1  GI:47525154
KEYWORDS    .
SOURCE      Pseudomonas syringae pv. maculicola
  ORGANISM  Pseudomonas syringae pv. maculicola
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 40110)
  AUTHORS   Stavrinides,J. and Guttman,D.S.
  TITLE     Nucleotide Sequence and Evolution of the Five-Plasmid Complement of
            the Phytopathogen Pseudomonas syringae pv. maculicola ES4326
  JOURNAL   J, Bacteriol. 186 (15), 5101-5115 (2004)
   PUBMED   15262947
REFERENCE   2  (bases 1 to 40110)
  AUTHORS   Stavrinides,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2004) Department of Botany, University of
            Toronto, 25 Willcocks Street, Toronto, ON M5S3B2, Canada
FEATURES             Location/Qualifiers
     source          1..40110
                     /organism="Pseudomonas syringae pv. maculicola"
                     /mol_type="genomic DNA"
                     /strain="ES4326"
                     /db_xref="taxon:59511"
                     /plasmid="pPMA4326B"
                     /note="pathovar: maculicola"
     gene            1..1302
                     /gene="repA"
                     /locus_tag="PMA4326B01"
     CDS             1..1302
                     /gene="repA"
                     /locus_tag="PMA4326B01"
                     /note="similar to Pseudomonas syringae pv. tomato DC3000
                     plasmid pDC3000A replication protein RepA encoded by
                     GenBank Accession Number AE016855"
                     /codon_start=1
                     /transl_table=11
                     /product="replication protein"
                     /protein_id="AAT35168.1"
                     /db_xref="GI:47525155"
                     /translation="MNHDNALSLSLAASHTANADPLASSTHLPPARFFEDGTALNRLL
                     LEAPYMARCSDDKTATRVRPREYALRYPYMQVNRPGMVSWLVFDLDHANALAWDDAGL
                     PAPNLMVRNRKSGHSQLFYAVPSVCTTENARAKPIQYMKAIYAAFAARLDADVDYHGG
                     PVAKTPGHPWWETTEFHSHVYELGELASAVELTVKPWATGPKLDQVSHSRHCILFEQL
                     RYFAYSIVNRERELGSFESFMRSLDAYAYNHNSFLKQGFSENLPLSSIRATVKSVGRW
                     TWDRYTGDRRCHRGAMQLDSSLSLTERQSLAARRTHELRHKATESKIRAACRQLQDQG
                     KALVRSAIATLAGVSASTVARYAHILTEVTKPATVSVLKAVPAANRDAPHGCREAVKS
                     PKKQGLADHDQGAPYGVHQISAVPEGPQAGEFLTIEHEDGS"
     gene            1451..1744
                     /gene="ssb"
                     /locus_tag="PMA4326B02"
     CDS             1451..1744
                     /gene="ssb"
                     /locus_tag="PMA4326B02"
                     /note="similar to Pseudomonas syringae pv. tomato DC3000
                     plasmid pDC3000A DNA-binding protein HU family encoded by
                     GenBank Accession Number AE016860"
                     /codon_start=1
                     /transl_table=11
                     /product="single-stranded binding protein"
                     /protein_id="AAT35169.1"
                     /db_xref="GI:47525156"
                     /translation="MALTKDQLVAGIAEAIDAPKTTARNALEQLGQIVADQLESGAEI
                     TLPGIGKLKVAERPARTGRNPSTGASIDIAAKKVVKFVSAKVLNDAINKGGVS"
     gene            1816..2016
                     /locus_tag="PMA4326B03"
     CDS             1816..2016
                     /locus_tag="PMA4326B03"
                     /note="similar to Escherichia coli O157:H7 EDL933 unknown
                     protein encoded by cryptic prophage CP-933P encoded by
                     GenBank Accession Number AE006461; possible
                     post-segregational killing system"
                     /codon_start=1
                     /transl_table=11
                     /product="stability determinant"
                     /protein_id="AAT35170.1"
                     /db_xref="GI:47525157"
                     /translation="MEKILMIDRSPIVSEFETEELEANYTAWLRAKVEASLADSRPAI
                     PHDEVERRMAERLARLRHRRAS"
     gene            2016..2309
                     /gene="parE"
                     /locus_tag="PMA4326B04"
     CDS             2016..2309
                     /gene="parE"
                     /locus_tag="PMA4326B04"
                     /note="similar to Escherichia coli O157:H7 hypothetical
                     protein encoded by GenBank Accession Number AP002557;
                     similar to plasmid stabilization system protein
                     (PFAM05016); possible post segregational killing system"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="stability determinant"
                     /protein_id="AAT35171.1"
                     /db_xref="GI:47525158"
                     /translation="MLPIFWLESADNDLAAIIEYIGLRDIAAAERLWQRLRGIVLPLS
                     EHPYLYAISDRVPGMREIVAHPNYLVFYRVTSTRIEVVNVVHARQEYPQTGLA"
     gene            2562..2975
                     /gene="PMA4326B05"
     CDS             2562..2975
                     /gene="PMA4326B05"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAT35172.1"
                     /db_xref="GI:47525159"
                     /translation="MVMLSLEQMLDALWSLEEAKLVAPLGSTWWSKAELNKQWGKPVV
                     TVSSESRLERLDMKFAAEHAKKATKATTSPSFLEMSEDETGAPYLLLEQTLLQCVYIA
                     EQLGAIPPLDEKWLARVIPPSLRQLSQLQKSTGLH"
     repeat_region   3001..5096
                     /note="inactive"
                     /insertion_seq="IS801"
     gene            3181..4741
                     /locus_tag="MFB01"
                     /note="similar to Pseudomonas syringae pv. maculicola M6
                     plasmid pFKN unknown encoded by GenBank Accession Number
                     AF359557; insertion caused disruption and frameshift in
                     the ORF; left and right borders still intact"
                     /pseudo
     gene            complement(5094..5256)
                     /locus_tag="MFB02"
                     /note="similar to Pseudomonas syringae pv. tomato str.
                     DC3000 ISPssy transposase encoded by GenBank Accession
                     Number AE016867"
                     /pseudo
     gene            5391..6809
                     /locus_tag="PMA4326B06"
     CDS             5391..6809
```

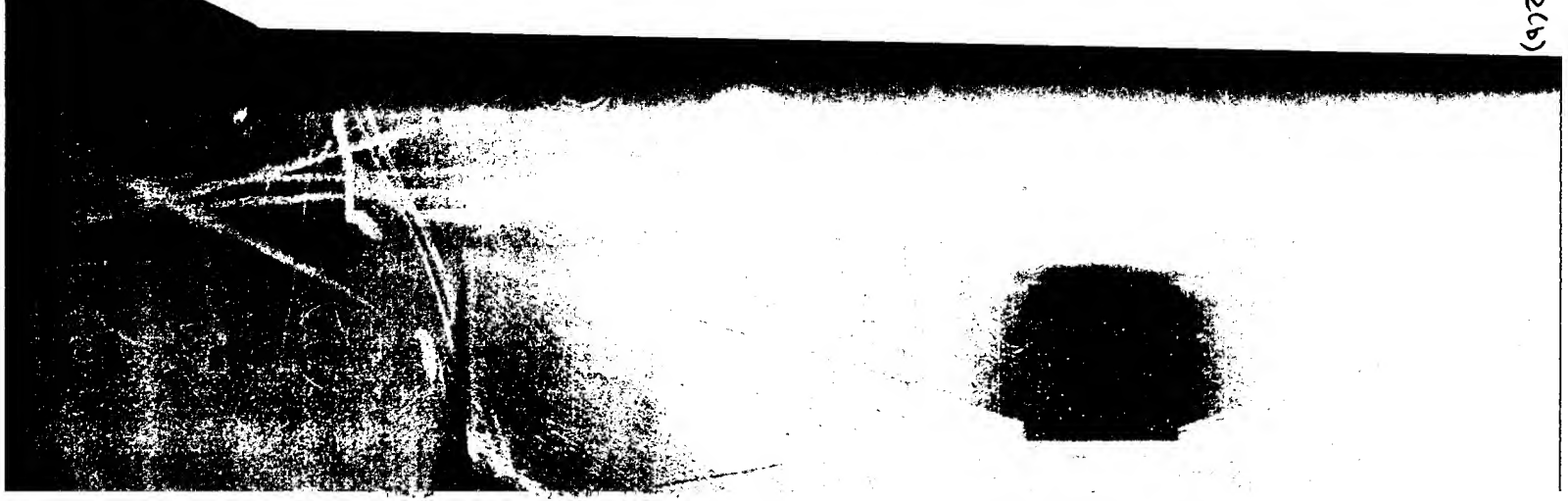
102(b)

```
                     /codon_start=1
                     /transl_table=11
                     /product="replication initiation protein"
                     /protein_id="AAC70076.1"
                     /db_xref="GI:3822122"
                     /translation="MTDLQQTYYRQVKNPNPVFTPREGAGTLKFCEKLMEKAVGFTSR
                     FDPAIHVAHARSKGLRRRMPPVLRRRAIDALLQGLCFHYDPLANRVQCSITTLAIECG
                     LATESAAGKLSITRATRALTFLSELGLITYQTEYDPLIGCYIPTDITFTPALFAALDV
                     SEDAVAAARRSRVEWENRLRKKQGLDTLGMDELIAKAWRFVRERFRSYQTELKSRGIK
                     RARARRDAGRERQDIVTLVKRQLTREISEGRFTANREAVKREVERRVKERMILSRNRN
                     YSRLATASP"
     misc_feature    3847. .4095
                     /standard_name="CIS"
                     /note="88 pct identical to locus ECCIS accession X12587,
                     required for cis-activation of oriR by the replication
                     initation protein"
     misc_feature    4072. .4080
                     /note="dnaA site; 100 pct identical (0 gaps) to locus
                     ECNR1REP at (1682. .1690) accession  X02302"
     rep_origin      4084. .4232
                     /standard_name="oriR"
                     /note="89 pct identical to oriR (1094. .1242); minimum
                     segment for replication of E. coli IncFII plasmid NR1
                     ECREPA1  X12776"
                     /direction=right
     promoter        4169. .4197
                     /note="predicted sigma 70 promoter; score of 56%"
     gene            4259. .4453
                     /gene="L7009"
     CDS             4259. .4453
                     /gene="L7009"
                     /note="95 pct identical to (0 gaps) 64 residues of a 128
                     aa protein REPA4 locus ECRSC1 accession V00351"
                     /codon_start=1

  Query Match             100.0%;  Score 26;  DB 1;  Length 92077;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAGGGGTTCCAAGCCGCAACTGACGA 26
              ||||||||||||||||||||||||||
Db       7262 AAGGGGTTCCAAGCCGCAACTGACGA 7237


RESULT 6
AX191727/c
LOCUS       AX191727               92077 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 9 from Patent WO0149775.
ACCESSION   AX191727
VERSION     AX191727.1  GI:15209896
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 9 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92077
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 92077;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAGGGGTTCCAAGCCGCAACTGACGA 26
              ||||||||||||||||||||||||||
Db       7262 AAGGGGTTCCAAGCCGCAACTGACGA 7237


RESULT 7
AB011549/c
LOCUS       AB011549               92721 bp    DNA     circular BCT 27-APR-1999
DEFINITION  Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION   AB011549
VERSION     AB011549.2  GI:4589740
KEYWORDS    ToxR-regulated lipoprotein; tagA.
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (sites)
  AUTHORS   Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
            Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
            Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
            and Shinagawa,H.
  TITLE     Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
            enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
            outbreak
  JOURNAL   DNA Res. 5 (1), 1-9 (1998)
  MEDLINE   98290540
   PUBMED   9628576
REFERENCE   2  (bases 1 to 92721)
  AUTHORS   Makino,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1998) Kozo Makino, Research Institute for
            Microbial Diseases, Osaka University, Molecular Microbiology;
            yamadaoka, 3-1, Suita, Osaka 562, Japan
            (E-mail:makino@bkns01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
            Fax:81-6-879-8320)
COMMENT     On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="O157:H7"
                     /sub_strain="RIMD 0509952"
                     /db_xref="taxon:562"
                     /plasmid="pO157"
                     /note="RIMD 0509952 is a strain of enterohemorrhagic E.
                     coli, EHEC O157:H7"
     gene            join(92527. .92721,1. .2502)
                     /gene="tagA"
     CDS             join(92527. .92721,1. .2502)
                     /gene="tagA"
                     /codon_start=1
                     /transl_table=11
                     /product="ToxR-regulated lipoprotein"
                     /protein_id="BAA31757.3"
                     /db_xref="GI:4666293"
                     /translation="MNTKMNERWRTPMKLKYLSCTILAPLAIGVFSATAADNNSAIYF
                     NTSQPINDLQGSLAAEVKFAQSQILPAHPKEGDSQPHLTSLRKSLLLVRPVKADDKTP
                     VQVEARDDNNKILGTLTLYPPSSLPDTIYHLDGVPEGGIDFTPHNGTKKIINTVAEVN
                     KLSDASGSSIHSHLTNNALVEIHTANGRWVRDIYLPQGPDLEGKMVRFVSSAGYSSTV
                     FYGDRKVTLSVGNTLLFKYVNGQWFRSGELENNRITYAQHIWSAELPAHWIVPGLNLV
                     IKQGNLSGRLNDIKIGAPGELLLHTIDIGMLTTPRDRFDFAKDKEAHREYFQTIPVSR
                     MIVNNYAPLHLKEVMLPTGELLTDMDPGNGGWHSGTMRQRIGKELVSHGIDNANYGLN
                     STAGLGENSHPYVVAQLAAHNSRGNYANGIQVHGGSGGGGIVTLDSTLGNEFSHEVGH
                     NYGLGHYVDGFKGSVHRSAENNNSTWGWDGDKKRFIPNFYPSQTNEKSCLNNQCQEPF
                     DGHKFGFDAMAGGSPFSAANRFTMYTPNSSAIIQRFFENKAVFDSRSSTGFSKWNADT
                     QEMEPYEHTIDRAEQITASVNELSESKMAELMAEYAVVKVHMWNGNWTRNIYIPTASA
                     DNRGSILTINHEAGYNSYLFINGDEKVVSQGYKKSFVSDGQFWKERDVVDTREARKPE
                     QFGVPVTTLVGYYDPEGTLSSYIYPAMYGAYGFTYSDDSQNLSDNDCQLQVDTKEGQL
                     RFRLANHRANNTVMNKFHINVPTESQPTQATLVCNNKILDTKSLTPAPEGLTYTVNGQ
                     ALPAKENEGCIVSVNSGKRYCLPVGQRSGYSLPDWIVGQEVYVDSGAKAKVLLSDWDN
                     LSYNRIGEFVGNVNPADMKKVKAWNGQYLDFSKPRSMRVVYK"
     gene            2589. .3464
                     /gene="etpC"
     CDS             2589. .3464
                     /gene="etpC"
```

```
                     /codon_start=1
                     /transl_table=11
                     /product="replication initiation protein"
                     /protein_id="AAC70076.1"
                     /db_xref="GI:3822122"
                     /translation="MTDLQQTYYRQVKNPNPVFTPREGAGTLKFCEKLMEKAVGFTSR
                     FDFAIHVAHARSKGLRRRMPPVLRRRAIDALLQGLCFHYDPLANRVQCSITTLAIECG
                     LATESAAGKLSITRATRALTFLSELGLITYQTEYDPLIGCYIPTDITFTPALFAALDV
                     SEDAVAAARRSRVEWENRLRKKQGLDTLGMDELIAKAWRFVRERFRSYQTELKSRGIK
                     RARARRDAGRERQDIVTLVKRQLTREISEGRFTANREAVKREVERRVKERMILSRNRN
                     YSRLATASP"
     misc_feature    3847. .4095
                     /standard_name="CIS"
                     /note="88 pct identical to locus ECCIS accession X12587,
                     required for cis-activation of oriR by the replication
                     initation protein"
     misc_feature    4072. .4080
                     /note="dnaA site; 100 pct identical (0 gaps) to locus
                     ECNR1REP at (1682. .1690) accession  X02302"
     rep_origin      4084. .4232
                     /standard_name="oriR"
                     /note="89 pct identical to oriR (1094. .1242); minimum
                     segment for replication of E. coli IncFII plasmid NR1
                     ECREPA1  X12776"
                     /direction=right
     promoter        4169. .4197
                     /note="predicted sigma 70 promoter; score of 56%"
     gene            4259. .4453
                     /gene="L7009"
     CDS             4259. .4453
                     /gene="L7009"
                     /note="95 pct identical to (0 gaps) 64 residues of a 128
                     aa protein REPA4 locus ECRSC1 accession V00351"
                     /codon_start=1

  Query Match             100.0%;  Score 26;  DB 1;  Length 92077;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGGTTCCAAGCCGCAACTGACG 26
              ||||||||||||||||||||||||||
Db       7263 TAAGGGGTTCCAAGCCGCAACTGACG 7238


RESULT 6
AX191727/c
LOCUS       AX191727               92077 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 9 from Patent WO0149775.
ACCESSION   AX191727
VERSION     AX191727.1  GI:15209896
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 9 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92077
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 92077;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAGGGGTTCCAAGCCGCAACTGACG 26
              ||||||||||||||||||||||||||
Db       7263 TAAGGGGTTCCAAGCCGCAACTGACG 7238


RESULT 7
AB011549/c
LOCUS       AB011549               92721 bp    DNA     circular BCT 27-APR-1999
DEFINITION  Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION   AB011549
VERSION     AB011549.2  GI:4589740
KEYWORDS    ToxR-regulated lipoprotein; tagA.
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (sites)
  AUTHORS   Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
            Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
            Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
            and Shinagawa,H.
  TITLE     Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
            enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
            outbreak
  JOURNAL   DNA Res. 5 (1), 1-9 (1998)
  MEDLINE   98290540
   PUBMED   9628576
REFERENCE   2  (bases 1 to 92721)
  AUTHORS   Makino,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1998) Kozo Makino, Research Institute for
            Microbial Diseases, Osaka University, Molecular Microbiology;
            yamadaoka, 3-1, Suita, Osaka 562, Japan
            (E-mail:makino@bkns01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
            Fax:81-6-879-8320)
COMMENT     On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="O157:H7"
                     /sub_strain="RIMD 0509952"
                     /db_xref="taxon:562"
                     /plasmid="pO157"
                     /note="RIMD 0509952 is a strain of enterohemorrhagic E.
                     coli, EHEC O157:H7"
     gene            join(92527. .92721,1. .2502)
                     /gene="tagA"
     CDS             join(92527. .92721,1. .2502)
                     /gene="tagA"
                     /codon_start=1
                     /transl_table=11
                     /product="ToxR-regulated lipoprotein"
                     /protein_id="BAA31757.3"
                     /db_xref="GI:4666293"
                     /translation="MNTKMNERWRTPMKLKYLSCTILAPLAIGVFSATAADNNSAIYF
                     NTSQPINDLQGSLAAEVKFAQSQILPAHPKEGDSQPHLTSLRKSLLLVRPVKADDKTP
                     VQVEARDDNNKILGTLTLYPPSSLPDTIYHLDGVPEGGIDFTPHNGTKKIINTVAEVN
                     KLSDASGSSIHSHLTNNALVEIHTANGRWVRDIYLPQGPDLEGKMVRFVSSAGYSSTV
                     FYGDRKVTLSVGNTLLFKYVNGQWFRSGELENNRITYAQHIWSAELPAHWIVPGLNLV
                     IKQGNLSGRLNDIKIGAPGELLLHTIDIGMLTTPRDRFDFAKDKEAHREYFQTIPVSR
                     MIVNNYAPLHLKEVMLPTGELLTDMDPGNGGWHSGTMRQRIGKELVSHGIDNANYGLN
                     STAGLGENSHPYVVAQLAAHNSRGNYANGIQVHGGSGGGGIVTLDSTLGNEFSHEVGH
                     NYGLGHYVDGFKGSVHRSAENNNSTWGWDGDKKRFIPNFYPSQTNEKSCLNNQCQEPF
                     DGHKFGFDAMAGGSPFSAANRFTMYTPNSSAIIQRFFENKAVFDSRSSTGFSKWNADT
                     QEMEPYEHTIDRAEQITASVNELSESKMAELMAEYAVVKVHMWNGNWTRNIYIPTASA
                     DNRGSILTINHEAGYNSYLFINGDEKVVSQGYKKSFVSDGQFWKERDVVDTREARKPE
                     QFGVPVTTLVGYYDPEGTLSSYIYPAMYGAYGFTYSDDSQNLSDNDCQLQVDTKEGQL
                     RFRLANHRANNTVMNKFHINVPTESQPTQATLVCNNKILDTKSLTPAPEGLTYTVNGQ
                     ALPAKENEGCIVSVNSGKRYCLPVGQRSGYSLPDWIVGQEVYVDSGAKAKVLLSDWDN
                     LSYNRIGEFVGNVNPADMKKVKAWNGQYLDFSKPRSMRVVYK"
     gene            2589. .3464
                     /gene="etpC"
     CDS             2589. .3464
                     /gene="etpC"
```

```
                     /codon_start=1
                     /transl_table=11
                     /product="replication initiation protein"
                     /protein_id="AAC70076.1"
                     /db_xref="GI:3822122"
                     /translation="MTDLQQTYYRQVKNPNPVFTPREGAGTLKFCEKLMEKAVGFTSR
                     FDFAIHVAHARSKGLRRRMPPVLRRRAIDALLQGLCFHYDPLANRVQCSITTLAIECG
                     LATESAAGKLSITRATRALTFLSELGLITYQTEYDPLIGCYIPTDITFTPALFAALDV
                     SEDAVAAARRSRVEWENRLRKKQGLDTLGMDELIAKAWRFVRERFRSYQTELKSRGIK
                     RARARRDAGRERQDIVTLVKRQLTREISEGRFTANREAVKREVERRVKERMILSRNRN
                     YSRLATASP"
     misc_feature    3847. .4095
                     /standard_name="CIS"
                     /note="88 pct identical to locus ECCIS accession X12587,
                     required for cis-activation of oriR by the replication
                     initation protein"
     misc_feature    4072. .4080
                     /note="dnaA site; 100 pct identical (0 gaps) to locus
                     ECNR1REP at (1682. .1690) accession  X02302"
     rep_origin      4084. .4232
                     /standard_name="oriR"
                     /note="89 pct identical to oriR (1094. .1242); minimum
                     segment for replication of E. coli IncFII plasmid NR1
                     ECREPA1  X12776"
                     /direction=right
     promoter        4169. .4197
                     /note="predicted sigma 70 promoter; score of 56%"
     gene            4259. .4453
                     /gene="L7009"
     CDS             4259. .4453
                     /gene="L7009"
                     /note="95 pct identical to (0 gaps) 64 residues of a 128
                     aa protein REPA4 locus ECRSC1 accession V00351"
                     /codon_start=1

  Query Match           100.0%;  Score 31;  DB 1;  Length 92077;
  Best Local Similarity  100.0%;  Pred. No. 0.017;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTCAACGGCATCGTCAGTTGCGGCTTGGAAC 31
              |||||||||||||||||||||||||||||||
Db       7227 CTCAACGGCATCGTCAGTTGCGGCTTGGAAC 7257


RESULT 6
AX191727
LOCUS       AX191727               92077 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 9 from Patent WO0149775.
ACCESSION   AX191727
VERSION     AX191727.1  GI:15209896
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 9 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92077
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match           100.0%;  Score 31;  DB 6;  Length 92077;
  Best Local Similarity  100.0%;  Pred. No. 0.017;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTCAACGGCATCGTCAGTTGCGGCTTGGAAC 31
              |||||||||||||||||||||||||||||||
```

```
Db       7227 CTCAACGGCATCGTCAGTTGCGGCTTGGA

RESULT 7
AB011549
LOCUS       AB011549                92721 bp
DEFINITION  Escherichia coli plasmid pO157
ACCESSION   AB011549
VERSION     AB011549.2  GI:4589740
KEYWORDS    ToxR-regulated lipoprotein; tag
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gamma
            Enterobacteriaceae; Escherichia
REFERENCE   1  (sites)
  AUTHORS   Makino,K., Ishii,K., Yasunaga,T
            Yutsudo,H.C., Kubota,Y., Yamaic
            Honda,T., Han,C.G., Ohtsubo,E.,
            and Shinagawa,H.
  TITLE     Complete nucleotide sequences o
            enterohemorrhagic Escherichia c
            outbreak
  JOURNAL   DNA Res. 5 (1), 1-9 (1998)
  MEDLINE   98290540
   PUBMED   9628576
REFERENCE   2  (bases 1 to 92721)
  AUTHORS   Makino,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1998) Kozo Ma
            Microbial Diseases, Osaka Unive
            yamadaoka, 3-1, Suita, Osaka 56
            (E-mail:makino@bkns01.biken.osa
            Fax:81-6-879-8320)
COMMENT     On Apr 20, 1999 this sequence ve
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia
                     /mol_type="genomic DNA"
                     /strain="O157:H7"
                     /sub_strain="RIMD 05099
                     /db_xref="taxon:562"
                     /plasmid="pO157"
                     /note="RIMD 0509952 is
                     coli, EHEC O157:H7"
     gene            join(92527. .92721,1. .
                     /gene="tagA"
     CDS             join(92527. .92721,1. .
                     /gene="tagA"
                     /codon_start=1
                     /transl_table=11
                     /product="ToxR-regulate
                     /protein_id="BAA31757.3
                     /db_xref="GI:4666293"
                     /translation="MNTKMNERW
                     NTSQPINDLQGSLAAEVKFAQSQ
                     VQVEARDDNNKILGTLTLYPPSS
                     KLSDASGSSIHSHLTNNALVEIH
                     FYGDRKVTLSVGNTLLFKYVNGQ
                     IKQGNLSGRLNDIKIGAPGELLL
                     MIVNNYAPLHLKEVMLPTGELLT
                     STAGLGENSHPYVVAQLAAHNSR
                     NYGLGHYVDGFKGSVHRSAENNN
                     DGHKFGFDAMAGGSPFSAANRFT
                     QEMEPYEHTIDRAEQITASVNEL
                     DNRGSILTINHEAGYNSYLFING
                     QFGVPVTTLVGYYDPEGTLSSYI
                     RFRLANHRANNTVMNKFHINVPT
                     ALPAKENEGCIVSVNSGKRYCLP
                     LSYNRIGEPVGNVNPADMKKVKA
     gene            2589. .3464
                     /gene="etpC"
     CDS             2589. .3464
                     /gene="etpC"
```



```
                TAQIAVQASLTGHLVLSTLHTNSAAGALSRLQDMGIPPFLLSTSLLAVLAQRLVRTLC
                PRCRQPCQVSTELAMDMDIPPETTIWQPAGCQHCSFTGYHGRTGIHELLLIDDRIRTA
                IYQGEGELGITRLAGSRYLTLRGDGRQKVLAGETSWEEVVRVTESRLQEEE"
gene            6939. .8162
                /gene="etpF"
CDS             6939. .8162
                /gene="etpF"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31761.1"
                /db_xref="GI:3337002"
                /translation="MALFHYQASDIHGRKRSGILEADSARHARQLLREQALIPVRLDE
                KQVHHKHSLRSILRFRPRGGSSAELALLTRQLATLVAASLPLEEALDALLRQSEKPRQ
                RNLIAAVRTKVLEGHSLAAAMGMFPGTFERLYCAMVAAGETSGRLDVVLSRLADYTEQ
                RQIMRNRLLQALLYPCVLTLVAVGVIAILLTAVVPKVVEQFIHMKQTLPLSTRVLMGA
                AEVSQTWGPWLLLAAALGGIAGRMILHQPSQRLAFHHLLLRLPVVGRISRGLNTARYA
                RTLSILNASAVPLLQAMHISGDVLSNDWARHQLATAAELVREGVSLHQALEQTSLFPP
                MMRHMIASGENSGELDSMLERAADNQDREFSTQMQLALGLFEPLLVVGMAGVVLFIVL
                AILQPLLQLNNMMNM"
gene            8193. .8627
                /gene="etpG"
CDS             8193. .8627
                /gene="etpG"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31762.1"
                /db_xref="GI:3337003"
                /translation="MRKQHQRGFTLLEIMVVIVILGVLASLVVPNLMGNKDKADRQKV
                MSDLVALESTLDMYRLDNNRYPTTEQGLRALVSKPTVQPEPRNYRQDGYIRRLPQDPW
                GGDYQLLNPGQYSDIDIFSPGPDGVPNTEDDIGNWTLGNAQP"
gene            8624. .9178
                /gene="etpH"
CDS             8624. .9178
                /gene="etpH"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31763.1"
                /db_xref="GI:3337004"
                /translation="MSQRGFTILEMMLILLLMGITAGLVLMSFPDSAQNHLQQQRERL
                QAQLDYALDRSQQDGLLMGIQVRFDGWKFKVLQRGTAESPPTLAEGGDIWQGYVWQTW
                QPRRAAMGGKLPDDVRLELQYLRGLQWMSEHDDGAEPDILLLPGGEVTPFRLLFRQVG
                EEAVVGLQVDENGLMTLFEGEVSL"
gene            9193. .9540
                /gene="etpI"
CDS             9193. .9540
                /gene="etpI"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31764.1"
                /db_xref="GI:3337005"
                /translation="MTLLEVIVALVVFALAGMALMQASTQQAAGIGRMEEKVLAGWLA
                DNQMVQLQLEKTWPENGWGEKTISFAGTEWYLRWGGPDSDVPQPRSLEVEVRRTKEET
                SALVSLRSSVVRE"
gene            9537. .10136
                /gene="etpJ"
CDS             9537. .10136
                /gene="etpJ"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31765.1"
                /db_xref="GI:3337006"
                /translation="MSQQRVKGFTLLEMLLALAVFAALSISAFQVLQGGIRAHELSRD
                KVQRLAELQRGISQMERDLTQMLPRHSRGNEGLLLAAPHLLKSDDWGISFTRNSWLNP
                AGMLPRPELQWVGYRLRQQKLERLSYFHVDHPSGVSPDVRVMLDGVHAFRLRFFVNGD
                WQARWDSTGILPQAVEVTLVMDDFAELPRLFLVSKETAE"
gene            10133. .11110
                /gene="etpK"
CDS             10133. .11110
                /gene="etpK"
                /codon_start=1
                /transl_table=11
                /product="Type II secretion pathway related protein"
                /protein_id="BAA31766.1"
                /db_xref="GI:3337007"
                /translation="MKLREQGVALLVVLLILSLMVTIAAVIAERNGRTFLRTV
                LQAKWDGYTAETIAKQILQRSRQESPRKTHLAQNWAQSERQFETRGGDVRGQI
                CFNLNAINYGVVDLTSIPYAARIFQQLLINLQVELLQARQVTAALRDWIDRDD
                GAEDEVYMGMEPPYLAANQPMQDVSELRLIRGIDARLYRKLLPYVCVLPTSDL
                TLLDSQAPLLAALFLTKPDSLPVTELLQRRPRTGWESVAAFLDAPPLKDIDTS
                LAVSSNYFLVRLHVRSGEHLFSQQTLMQWREERFRIIQRQYGLTMREVP"
gene            11473. .12321

  Query Match             100.0%;  Score 16;  DB 1;  Length 92721;
  Best Local Similarity   100.0%;  Pred. No. 5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps

Qy          1 CGGCATCGTCAGTTGC 16
              ||||||||||||||||
Db      76747 CGGCATCGTCAGTTGC 76762


RESULT 16
AX191725
LOCUS       AX191725               92721 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 7 from Patent WO0149775.
ACCESSION   AX191725
VERSION     AX191725.1  GI:15209894
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 7 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 16;  DB 6;  Length 92721;
  Best Local Similarity   100.0%;  Pred. No. 5e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGCATCGTCAGTTGC 16
              ||||||||||||||||
Db      76747 CGGCATCGTCAGTTGC 76762


RESULT 17
UBA319838
LOCUS       UBA319838              536 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Uncultured bacterium SB-53-TW partial 16S rRNA gene.
ACCESSION   AJ319838
VERSION     AJ319838.1  GI:21489345
KEYWORDS    16S ribosomal RNA; 16S rRNA gene.
SOURCE      uncultured bacterium
  ORGANISM  uncultured bacterium
            Bacteria; environmental samples.
REFERENCE   1
  AUTHORS   Schaefer,H., Abbas,B., Witte,H. and Muyzer,G.
  TITLE     Genetic diversity of 'satellite' bacteria present in cultures of
            marine diatoms
  JOURNAL   FEMS Microbiol. Ecol. 42 (1), 25-35 (2002)
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Schaefer,H.
  TITLE     Direct Submission
```

```
LOCUS       AF074613               92077 bp    DNA     circular BCT 04-NOV-1998
DEFINITION  Escherichia coli O157:H7 plasmid pO157, complete sequence.
ACCESSION   AF074613
VERSION     AF074613.1  GI:3822114
KEYWORDS    .
SOURCE      Escherichia coli O157:H7
  ORGANISM  Escherichia coli O157:H7
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 92077)
  AUTHORS   Burland,V., Shao,Y., Perna,N.T., Plunkett,G., Sofia,H.J. and
            Blattner,F.R.
  TITLE     The complete DNA sequence and analysis of the large virulence
            plasmid of Escherichia coli O157:H7
  JOURNAL   Nucleic Acids Res. 26 (18), 4196-4204 (1998)
  MEDLINE   98391744
   PUBMED   9722640
REFERENCE   2  (bases 1 to 92077)
  AUTHORS   Burland,V., Shao,Y., Perna,N.T., Plunkett,G. III, Sofia,H.J. and
            Blattner,F.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445
            Henry Mall, Madison, WI 53706, USA
FEATURES             Location/Qualifiers
     source          1. .92077
                     /organism="Escherichia coli O157:H7"
                     /mol_type="genomic DNA"
                     /strain="EDL933"
                     /serotype="O157:H7"
                     /db_xref="taxon:83334"
                     /lab_host="Escherichia coli C600"
                     /plasmid="pO157"
     gene            1. .561
                     /gene="finO"
                     /note="synonym: L7001"
     CDS             1. .561
                     /gene="finO"
                     /note="97 pct identical amino acid sequence and equal
                     length to FIO1_ECOLI SW:  P22707"
                     /codon_start=1
                     /transl_table=11
                     /product="fertility inhibition protein (conjugal transfer
                     repressor)"
                     /protein_id="AAC70069.1"
                     /db_xref="GI:3822115"
                     /translation="MAEQKRPVLTLKRKTEGETPVRSRKTIINVTTPPKWKVKKQKLA
                     EKAAREAELAAKKAQARQALSIYLNLPTQDEAVNTLKPWWPGLFDGDTPRLLACGIRD
                     VLLEDVAQRNIPLSHKKLRRALKAITRSESYLCAMKAGACRYDTEGYVTEHISQEEEA
                     YAAARLDKIRRQNRIKAELQAVLDEK"
     promoter        653. .681
                     /note="predicted sigma 70 promoter; score of 52%"
     gene            698. .949
                     /gene="L7002"
     CDS             698. .949
                     /gene="L7002"
                     /note="54 pct identical (1 gap) to 23 residues of an
                     approx. 269 aa protein EXOK_RHIME SW:  P33693"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAC70070.1"
                     /db_xref="GI:3822116"
                     /translation="MDSETVHGTVRSGVTSVPEAGPLFWKSVDAGWKRQKHGDGLPVL
                     HPGLTGSSLPEKGLNTATGAAEEGGNEKSSLHYSRDSEG"
     promoter        1025. .1052
                     /note="predicted sigma 70 promoter; score of 62%"
     gene            1151. .1612
                     /gene="L7003"
     CDS             1151. .1612
                     /gene="L7003"
                     /note="98 pct identical and equal length to YFI3_ECOLI SW:
                     Q99342"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein 15.6 kDa protein in finO 3'
                     region precursor"
                     /protein_id="AAC70071.1"
                     /db_xref="GI:3822117"
                     /translation="MRKYIPLVLFIFSWPVLCADIHGRVVRVLDGDTIEVMDSRKAVR
                     IRLVNIDAPEKKQDYGRWSTDMMKSLVAGKTVTVTYFQRDRYGRMLGQVYAPDGMNVN
                     QFMVRAGAAWVYEQYNTDPVLPVLQNEARQQKRGLWSDADPVPPWIWRHRK"
     promoter        1390. .1419
                     /note="predicted sigma 70 promoter; score of 56%"
     gene            1658. .1867
                     /gene="L7004"
     CDS             1658. .1867
                     /gene="L7004"
                     /note="55 pct identical (0 gaps) to 66 residues of an
                     approx. 72 aa protein HHA_ECOLI SW:  P23870"
                     /codon_start=1
                     /transl_table=11
                     /product="putative hemolysin expression modulating
                     protein"
                     /protein_id="AAC70072.1"
                     /db_xref="GI:3822118"
                     /translation="MEKTKQEWLYQLRRCSSVNTLEKIIHKNRDSLSTSERESFNSAA
                     DHRLAELITGKLYDRIPKEIWKYVR"
     promoter        1899. .1927
                     /note="predicted sigma 70 promoter; score of 56%"
     gene            1905. .2243
                     /gene="L7005"
     CDS             1905. .2243
                     /gene="L7005"
                     /note="98 pct identical (0 gaps) to 57 residues of an
                     approx. 200 aa protein; plasmid R100 miniplasmid  pSM1 ORF
                     4, TRSW:  Q52340"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAC70073.1"
                     /db_xref="GI:3822119"
                     /translation="MKLIIFILIVLIIAALLIRIILRSVNQHSPLLMQLHAAGIRTGD
                     AERILSGGEYDASHLRPERRRRDILLEEVLRQAGIPLLRSHDARKLLQMTGEWLNTTG
                     AAQQSPEHRS"
     misc_feature    complement(2240. .2327)
                     /note="initiation site; Escherichia coli plasmid R100 ssiB
                     gene; 99 pct  identical to ssiB locus R10SSIB accession
                     D90185"
     -35_signal      2416. .2421
                     /note="100 pct identical (0 gaps) to the -35  region at
                     23. .28 locus  ECNR1REP accession X02302"
     -10_signal      2439. .2445
                     /note="100 pct identical (0 gaps) to the -10  region at
                     46. .52 locus  ECNR1REP accession X02302"
     RBS             2473. .2478
                     /note="100 pct identical (0 gaps) to RBS at 80. .85 locus
                     ECNR1REP accession X02302"
     gene            2483. .2737
                     /gene="cpb2"
                     /note="synonym: L7006"
     CDS             2483. .2737
                     /gene="cpb2"
                     /note="98 pct identical amino acid sequence and equal
                     length to CPB2_ECOLI SW:  P03847"
                     /codon_start=1
                     /transl_table=11
                     /product="CopB protein (RepA2 protein)"
                     /protein_id="AAC70074.1"
                     /db_xref="GI:3822120"
                     /translation="MSQTENAVTSSSGAKRAYRKGNPLSDAEKQRLSVARKRASFKEV
                     KVFLEPKYKAMLMQMCHEDGLTQAEVLTALIKSEAQKRCV"
     promoter        2754. .2782
                     /note="predicted sigma 70 promoter; score of 69%"
     gene            2973. .3047
                     /gene="L7007"
     CDS             2973. .3047
```

```
                     /gene="L7007"
                     /note="91 pct identical (0 gaps) to 24 residues of an
                     approx. 24 aa protein uORF  P30REPPIC, accession M16167,
                     translationally coupled to replication initiation protein"
                     /codon_start=1
                     /transl_table=11
                     /product="replication initiation protein"
                     /protein_id="AAC70075.1"
                     /db_xref="GI:3822121"
                     /translation="MLGKVQDFFLCSLLLRIVSAGWCD"
     RBS             3032. .3035
                     /note="100 pct identical (0 gaps) to RBS at 641. .644 locus
                      ECNR1REP accession X02302"
     gene            3040. .3897
                     /gene="rep2"
                     /note="synonym: L7008"
     CDS             3040. .3897
                     /gene="rep2"
                     /note="96 pct identical amino acid sequence and equal
                     length to REP2_ECOLI SW:  P03066"
                     /codon_start=1
                     /transl_table=11
                     /product="replication initiation protein"
                     /protein_id="AAC70076.1"
                     /db_xref="GI:3822122"
                     /translation="MTDLQQTYYRQVKNPNPVFTPREGAGTLKFCEKLMEKAVGFTSR
                     FDFAIHVAHARSKGLRRRMPPVLRRRAIDALLQGLCFHYDPLANRVQCSITTLAIECG
                     LATESAAGKLSITRATRALTFLSELGLITYQTEYDPLIGCYIPTDITFTPALFAALDV
                     SEDAVAAARRSRVEWENRLRKKQGLDTLGMDELIAKAWRFVRERFRSYQTELKSRGIK
                     RARARRDAGRERQDIVTLVKRQLTREISEGRFTANREAVKREVERRVKERMILSRNRN
                     YSRLATASP"
     misc_feature    3847. .4095
                     /standard_name="CIS"
                     /note="88 pct identical to locus ECCIS accession X12587,
                     required for cis-activation of oriR by the replication
                     initation protein"
     misc_feature    4072. .4080
                     /note="dnaA site; 100 pct identical (0 gaps) to locus
                     ECNR1REP at (1682. .1690) accession  X02302"
     rep_origin      4084. .4232
                     /standard_name="oriR"
                     /note="89 pct identical to oriR (1094. .1242); minimum
                     segment for replication of E. coli IncFII plasmid NR1
                     ECREPA1  X12776"
                     /direction=right
     promoter        4169. .4197
                     /note="predicted sigma 70 promoter; score of 56%"
     gene            4259. .4453
                     /gene="L7009"
     CDS             4259. .4453
                     /gene="L7009"
                     /note="95 pct identical to (0 gaps) 64 residues of a 128
                     aa protein REPA4 locus ECRSC1 accession V00351"
                     /codon_start=1

  Query Match             100.0%;  Score 18;  DB 1;  Length 92077;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCATCGTCAGTTGCG 18
              ||||||||||||||||||
Db       7231 ACGGCATCGTCAGTTGCG 7248


RESULT 10
AX191727
LOCUS       AX191727               92077 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 9 from Patent WO0149775.
ACCESSION   AX191727
VERSION     AX191727.1  GI:15209896
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 9 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .92077
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 18;  DB 6;  Length 92077;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCATCGTCAGTTGCG 18
              ||||||||||||||||||
Db       7231 ACGGCATCGTCAGTTGCG 7248


RESULT 11
AB011549
LOCUS       AB011549               92721 bp    DNA     circular BCT 27-APR-1999
DEFINITION  Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION   AB011549
VERSION     AB011549.2  GI:4589740
KEYWORDS    ToxR-regulated lipoprotein; tagA.
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (sites)
  AUTHORS   Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
            Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
            Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
            and Shinagawa,H.
  TITLE     Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
            enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
            outbreak
  JOURNAL   DNA Res. 5 (1), 1-9 (1998)
  MEDLINE   98290540
   PUBMED   9628576
REFERENCE   2  (bases 1 to 92721)
  AUTHORS   Makino,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1998) Kozo Makino, Research Institute for
            Microbial Diseases, Osaka University, Molecular Microbiology;
            yamadaoka, 3-1, Suita, Osaka 562, Japan
            (E-mail:makino@bkns01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
            Fax:81-6-879-8320)
COMMENT     On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES             Location/Qualifiers
     source          1. .92721
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /strain="O157:H7"
                     /sub_strain="RIMD 0509952"
                     /db_xref="taxon:562"
                     /plasmid="pO157"
                     /note="RIMD 0509952 is a strain of enterohemorrhagic E.
                     coli, EHEC O157:H7"
     gene            join(92527. .92721,1. .2502)
                     /gene="tagA"
     CDS             join(92527. .92721,1. .2502)
                     /gene="tagA"
                     /codon_start=1
                     /transl_table=11
                     /product="ToxR-regulated lipoprotein"
                     /protein_id="BAA31757.3"
                     /db_xref="GI:4666293"
```